

Univerzita Karlova  
Filozofická fakulta  
Ústav českého jazyka a teorie komunikace

Diplomová práce

**Vzájemné hodnocení ve výuce češtiny jako cizího jazyka**

Peer Assessment in Teaching Czech as a Foreign Language

Bc. Anna Stuchlá

### **Čestné prohlášení:**

Prohlašuji, že jsem diplomovou práci vypracovala samostatně, že jsem řádně citovala všechny použité prameny a literaturu a že práce nebyla využita v rámci jiného vysokoškolského studia či k získání jiného nebo stejného titulu.

V Praze dne 10. května 2021

.....

Bc. Anna Stuchlá

**Poděkování:**

Na tomto místě bych ráda poděkovala zejména Mgr. Janě Šindlerové, Ph.D. za vstřícné a empatické vedení a za veškerou odbornou podporu, kterou mi během psaní poskytla.

**Klíčová slova:**

formativní hodnocení, vzájemné hodnocení, pedagogické hodnocení, čeština jako druhý/cizí jazyk, výuka psaní, úroveň B1, zkouška z češtiny pro udělení státního občanství

**Key words:**

formative assessment, peer-to-peer assessment, czech as a second/foreign language, teaching writing skills, B1 level, Czech language exam for candidates for Czech citizenship

## **Abstrakt:**

Diplomová práce si klade za cíl (i) ověřit vhodnost využití vzájemného formativního hodnocení ve výuce psaní v kurzech češtiny jako cizího jazyka na úrovni B1 a (ii) zformulovat možná doporučení pro lektory češtiny jako cizího jazyka, kteří by uvažovali o případném využití tohoto postupu ve svých hodinách.

V teoretické části se text věnuje jednak popisu jednotlivých typů a forem hodnocení, jednak shrnuje výsledky několika studií, které v této oblasti vznikly zejména ve výzkumu didaktiky angličtiny. Praktická část práce pak popisuje přípravu, průběh a výsledky experimentu, který byl podniknut se skupinou studentů jako intenzivní víkendový online kurz psaní. Experimentu se zúčastnili studenti připravující se k písemné části zkoušky z češtiny na úrovni B1.

Výzkum kvality komentářů od studentů a míry realizace jejich návrhů úprav ukázal omezenou efektivitu vzájemného hodnocení ve výuce studentů na dané úrovni. Na základě získaných dat se však podařilo identifikovat některé nedostatky ve způsobu, jímž byla metoda v experimentálním kurzu využita, a zformulovat doporučení, jejichž dodržení by tyto nedostatky mohlo odstranit. To by v důsledku mohlo vést k vhodnější implementaci vzájemného hodnocení v kurzech psaní na úrovni B1.

## **Abstract:**

The aims of this Master's thesis are (i) to examine the suitability of using peer-to-peer assessment in teaching writing skills in Czech B1-level courses and (ii) to suggest recommendations for tutors who are considering implementing the method in their Czech classes.

The theoretical part is concerned with the description of particular types and forms of assessment and it presents an overview of several studies, focusing on those conducted in the field of ESL teaching. The empirical part focuses on describing the design and execution of the experiment as well as on interpreting its results. The experiment was performed as an intensive online writing course in Czech; the participants were students preparing for a B1-level exam.

The investigation aimed at the quality of students' comments and at the implementation degree of their suggestions shows limited effectivity of using peer-to-peer assessment in teaching Czech on B1-level. However, based on the data, the study managed to identify several shortcomings in the way the peer-to-peer assessment method was used and it suggested recommendations for their elimination. If the tutors decide to follow these suggestions this could lead to a more successful use of the investigated method in teaching.

# Obsah

<b>1 Úvod</b>	<b>9</b>
<b>2 Hodnocení jako pedagogický koncept</b>	<b>10</b>
2.1 Funkce hodnocení	10
2.2 Typologie hodnocení	10
2.2.1 Assessment vs. evaluation	11
2.2.2 Bezděčné vs. záměrné hodnocení	11
2.2.3 Holistické vs. analytické hodnocení	11
2.2.4 Sumativní vs. formativní hodnocení	11
2.2.5 Normativní vs. kritériální hodnocení	11
2.2.6 Pedagogické a vzájemné hodnocení, sebehodnocení	12
2.3 Specifika hodnocení slohových prací	12
<b>3 Formativní hodnocení</b>	<b>13</b>
3.1 Vymezení pojmu	13
3.2 Vývoj formativního hodnocení	13
3.2.1 Mastery learning	14
3.2.2 Assessment for learning	14
3.3 Zásady pro implementaci	15
3.3.1 Efektivní plánování	15
3.3.2 Zaměření na proces učení	16
3.3.3 Centrální postavení ve výukové praxi	16
3.3.4 Klíčová dovednost učitele	16
3.3.5 Citlivost a konstruktivnost	16
3.3.6 Podpora motivace	16
3.3.7 Porozumění cílům a kritériím	17
3.3.8 Napomáhání seberochoji	17
3.3.9 Rozvíjení dovednosti sebehodnocení	17
3.3.10 Rozpoznávání všech výsledků vzdělání	17
3.4 Formativní hodnocení a výuka cizích jazyků	17
3.4.1 Výzkumy formativního hodnocení v jazykové výuce	17

3.4.2	Hodnocení ve výuce češtiny jako cizího jazyka	18
<b>4</b>	<b>Vzájemné hodnocení</b>	<b>20</b>
4.1	Dosavadní výzkumy vzájemného hodnocení	20
4.1.1	Známkování	20
4.1.2	Vnímání vzájemného hodnocení ze strany studentů	21
4.1.3	Vzájemné hodnocení v jazykové výuce	22
4.2	Zásady pro implementaci vzájemného hodnocení	23
4.2.1	Jazyková úroveň	24
4.2.2	Sociální a věková skupina	24
4.3	Online nástroje pro vzájemné hodnocení	24
<b>5</b>	<b>Příprava výzkumu</b>	<b>26</b>
5.1	Cíle výzkumu a jeho základní charakteristika	26
5.2	Výzkumné otázky	26
5.2.1	Otázky týkající se objektivní efektivity metody	27
5.2.2	Otázky týkající se subjektivního hodnocení přínosu ze strany studentů	27
5.3	Výzkumné hypotézy	27
5.4	Východiska experimentu	28
5.4.1	Úroveň B1	28
5.4.2	Zkouška z češtiny pro udělení státního občanství	29
5.4.3	Certifikovaná zkouška z češtiny CCE-B1	32
5.4.4	Kritéria hodnocení písemných prací	33
5.4.5	Reflexe zkoušek v podobě experimentu	34
5.5	Organizace a popis experimentu	34
5.5.1	První fáze: Stanovení kritérií	35
5.5.2	Druhá fáze: Psaní úlohy	35
5.5.3	Třetí fáze: Hodnocení	36
5.5.4	Čtvrtá fáze: Diskuse nad obdrženým hodnocením	37
5.5.5	Pátá fáze: Druhá verze textu	37
5.6	Dotazník	38

<b>6</b>	<b>Vyhodnocení výzkumu</b>	<b>39</b>
6.1	Respondenti	39
6.2	Nasbíraná data	39
6.2.1	Zpracování dat z textů	40
6.2.2	Zpracování dotazníků	42
6.3	Výsledky ve vztahu k výzkumným otázkám	42
6.3.1	Kvalita komentářů u vzájemného hodnocení	42
6.3.2	Implementace úprav	48
6.3.3	Vnímání vzájemného hodnocení studenty	53
6.4	Diskuse	57
6.5	Doporučení pro lektory	58
6.5.1	Předběžné vyhodnocení vhodnosti využití v dané skupině	58
6.5.2	Postupná implementace a nácvik	59
6.5.3	Doplnění o zpětnou vazbu od lektora	59
<b>7</b>	<b>Shrnutí a závěr</b>	<b>60</b>
<b>8</b>	<b>Seznam použité literatury</b>	<b>62</b>
<b>9</b>	<b>Seznam příloh</b>	<b>66</b>



# 1 Úvod

Psaní je vedle poslechu, mluvení a čtení jednou ze čtyř klíčových oblastí, které potřebují studenti zvládnout, pokud si chtějí cizí jazyk zcela osvojit. Přesto můžeme v praxi mnohdy pozorovat, že je v kurzech češtiny pro cizince rozvíjení písemného projevu studentů odsunuto z různých důvodů „na druhou kolej“ a studenti se s ním často setkávají pouze v podobě písemných domácích úkolů. Potřebu systematictější práce mnoho z nich obvykle pocítí až ve chvíli, kdy se připravují na zkoušky, jejichž součástí jsou i písemné úlohy. V kurzech zaměřených na přípravu ke zkoušce na úrovni B1, kterou potřebují zvládnout zájemci o české státní občanství, může být proto na lektora kladen poměrně velký tlak ohledně intenzivního nácviku psacích dovedností. Neustálé opravování písemných prací pro něj pak samozřejmě představuje velkou časovou zátěž.

Vzájemné hodnocení ve výuce jazyka může představovat v rámci formativního hodnocení alternativu a doplněk k hodnocení pedagogem. Vedle možné úspory času lektora navíc jeho využití potenciálně nabízí i další výhody, jako je např. rozvíjení kooperativního učení a metakognitivních kompetencí studentů (Cheng – Warren, 2005, s. 94). Přesto se mu ve výzkumu didaktiky češtiny jako cizího jazyka dostává pozornosti pouze okrajově. Tato práce se proto věnuje vzájemnému hodnocení ve výuce psaní a dává si za cíl (i) ověřit vhodnost využití vzájemného formativního hodnocení ve výuce psaní v kurzech češtiny jako cizího jazyka na úrovni B1 a (ii) zformulovat možná doporučení pro lektory češtiny jako cizího jazyka, kteří uvažují o případném využití tohoto postupu ve svých hodinách.

Ve druhé, třetí a čtvrté kapitole diplomové práce shrnuji relevantní dostupnou literaturu k tématu. Druhá kapitola se obecně věnuje úloze, typům a formám hodnocení v pedagogice. Ve třetí kapitole představuji koncept formativního hodnocení, a to včetně doporučení pro jeho implementaci do výuky cizího jazyka. Čtvrtá kapitola se pak blíže zaměřuje na vzájemné (formativní) hodnocení v kontrastu k hodnocení pedagogickému. Prostor je věnován možným výhodám, jež tyto postupy nabízejí, a také zásadám pro jejich uplatnění ve výuce.

V páté a šesté kapitole této práce popisují výzkum využitelnosti vzájemného hodnocení ve výuce studentů, kteří se připravují k písemné části zkoušky na úrovni B1. Pátá kapitola se věnuje výzkumným otázkám, hypotézám a přípravě experimentu. V šesté kapitole jsou představeny výsledky a na jejich základě jsou zformulována doporučení pro praktické využití vzájemného hodnocení ve výuce. Vycházím přitom jednak z dat získaných v samotném experimentu, jež jsem uskutečnila v rámci intenzivního víkendového online kurzu psaní, jednak z odpovědí z dotazníku, který měli respondenti za úkol vyplnit po skončení tohoto kurzu. V sedmé kapitole práce shrnuji nejdůležitější závěry a nastiňuji možnosti dalšího výzkumu.

## **2 Hodnocení jako pedagogický koncept**

Podle Slavíka (1999, s. 15–16) můžeme hodnocení ve škole na obecné rovině definovat jako porovnávání dvou objektů, při němž se snažíme rozlišit nebo vybrat to, co je lepší a to, co je horší, a případně také najít cestu k vylepšení nebo nápravě toho horšího. Objektem hodnocení je pak obvykle produkt studenta, který je srovnáván buďto s jiným srovnatelným objektem nebo s nějakým ideálním vzorem či normou.

Podstatu školního hodnocení lze definovat jako „zjištění kompetencí – vědomostí, dovedností a postojů – žáka a formulace hodnotících soudů na základě porovnání skutečného stavu se stavem předpokládaným, formulovaným jako cíle výuky.“ (Kalhous – Obst, 2002, s. 404).

Hodnocení představuje klíčovou aktivitu jak v prostředí školy, tak ve vzdělávání dospělých. Samotnými učiteli je pak podle některých studií (např. Bendl, 1997) subjektivně vnímáno jako jedna z náročnějších částí jejich práce.

### **2.1 Funkce hodnocení**

Hodnocení má ve výuce směrem ke studentovi hned několik funkcí. Jejich přehled nabízí např. Slavík (1999, s. 17–18), který vymezuje ve vztahu přímých aktérů ve škole funkci motivační, poznávací a konativní.

Motivační funkce je spojena s citovým prožíváním a „zasahuje do intimní sféry“ (tamtéž, s. 17) hodnoceného i hodnotitele. V praxi bývá realizována osobními otázkami.

Poznávací funkce souvisí spíše s intelektuální dimenzí a týká se předmětu hodnocení. Bývá realizována analytickými otázkami.

Konativní funkce „směřuje k aktivnímu působení“ (tamtéž, s. 18). Je spojena s vůlí a bývá realizována „otázkou po aktivitě“. Hodnoceného by tyto otázky měly vést k tomu, aby pátral po tom, co lze udělat, aby se jeho výkon zlepšil.

### **2.2 Typologie hodnocení**

Dnešní pedagogika může na základě různých aspektů dojít k různému rozdělení hodnocení na jednotlivé typy. Slavík (1999, s. 34–41) např. nabízí následující klasifikaci: (i) „assessment“ vs. „evaluation“, (ii) bezděčné vs. záměrné hodnocení, (iii) holistické vs. analytické hodnocení, (iv) sumativní vs. formativní hodnocení a (v) normativní vs. kritériální hodnocení. Vedle toho můžeme hodnocení také rozdělit podle formy zapojení aktérů na (i) pedagogické hodnocení, (ii) vzájemné hodnocení a (iii) sebehodnocení.

### **2.2.1 Assessment vs. evaluation**

Pod vlivem angličtiny se někdy v odborné literatuře rozlišuje mezi hodnocením („assessment“) a evaluací (Slavík, 1999, s. 35). Pojem evaluace se postupně ustálil pro strategické vyhodnocování kvality „různých pojetí vzdělávání a různých typů edukativních aktivit a jejich výsledků“ ve vzdělávacím systému, zatímco pojem hodnocení ve smyslu „assessment“ se využívá pro vnitřní hodnocení ve třídě, jehož náplní je obecně „diagnostický sběr a posuzování hodnotících údajů“, který umožňuje „bezprostřední korekční zásahy a průběžné plánování dalšího postupu vzhledem ke stanoveným cílům“ (tamtéž, s. 36–37). Právě hodnocení ve smyslu „assessment“ se tedy věnuje i tato diplomová práce.

### **2.2.2 Bezděčné vs. záměrné hodnocení**

Mezi bezděčným a záměrným hodnocením není vždy stanovena pevná hranice. U záměrného hodnocení je ale zřetelně přítomná kontrola vědomí a vůle, a takové hodnocení lze tedy analyzovat a formalizovat (Slavík, 1999, s. 35), zatímco bezděčné hodnocení je spontánní a citové (tamtéž, s. 34).

### **2.2.3 Holistické vs. analytické hodnocení**

Zatímco v případě holistického hodnocení jde o povšechné hodnocení výkonu jako celku, analytické hodnocení se blíže soustředí na posouzení jeho dílčích částí (Slavík 1999, s. 34–35).

### **2.2.4 Sumativní vs. formativní hodnocení**

Cílem sumativního (neboli také finálního) hodnocení je získat „konečný celkový přehled o dosahovaných výkonech nebo kvalitativně roztrždit celý posuzovaný soubor“ (Slavík, 1999, s. 37). Formativní (korektivní, pracovní, zpětnovazební) hodnocení se naproti tomu soustředí na proces učení a poskytuje studentům zpětnou vazbu v jeho průběhu, tedy v momentu, kdy se jejich výkon může na jejím základě ještě zlepšit (tamtéž, s. 38). Jeho implementace má podle Slavíka (tamtéž, s. 39) „větší šanci být ze stany posuzovaného chápán[a] jako pomoc, prostředek k vlastnímu zlepšení a k sebepoznání, nikoli jako rozsudek nebo odsudek“, a má tedy pozitivní vliv na motivaci člověka, a to dokonce i bez ohledu na jeho výsledek.

### **2.2.5 Normativní vs. kritériální hodnocení**

Rozlišní normativního a kritériálního hodnocení je založeno na dvou možných měřících, na základě kterých je výkon posuzován. U normativního hodnocení jde o normu stanovenou vzhledem k určité referenční skupině studentů, zatímco u kritériálního hodnocení vztahujeme výkon k předem daným kritériím, a jde tedy o posuzování absolutního výkonu, kde nedochází

k poměrování s ostatními studenty, ale záleží na splnění předem stanovených podmínek úspěšnosti (Slavík, 1999, s. 40).

### **2.2.6 Pedagogické a vzájemné hodnocení, sebehodnocení**

Na základě zapojení aktérů (tzn. kdo koho hodnotí) můžeme hodnocení rozdělit také na pedagogické (autoritativní), vzájemné (vrstevnické) a na sebehodnocení.

Pedagogické hodnocení, jež je ve výuce asi nejobvyklejší, je profesionální zpětná vazba, kterou poskytuje učitel studentům. Vrstevnické hodnocení a sebehodnocení naproti tomu přesouvají zodpovědnost za zpětnou vazbu na studenty, každé však jiným způsobem. Zatímco vrstevnické hodnocení pracuje s aktivitami, během nichž si studenti poskytují zpětnou vazbu vzájemně se spolužáky, sebehodnocení představuje proces, během něž studenti sami analyzují a vyhodnocují vlastní výkon (Falchikov, 2000, s. 287).

V metaanalýze studií zaměřených na formativní hodnocení, kterou podnikly Floréz<sup>1</sup> a Sammons (2013), se uvádí, že vzájemné hodnocení i sebehodnocení může rozvíjet autonomii a metakognitivní dovednosti studentů a napomáhat kooperativním aktivitám ve třídě (tamtéž, s. 10).

## **2.3 Specifika hodnocení slohových prací**

Specifickou oblast ve výuce mateřského i cizího jazyka představuje hodnocení slohových prací. Náročnost takového úkolu pro učitele nastiňuje např. Čechová (1998, s. 245). Ta ji vysvětluje zejména komplexností hodnocení, které musí vycházet z mnoha různých dílčích složek, jež ovšem mohou být v písemném projevu na velmi odlišné úrovni. Tyto dílčí složky může představovat např. gramatická správnost použitých prostředků, vhodnost a šíře použité slovní zásoby, naplnění komunikačních cílů vzhledem k typu textu atd.

Standardizované zkoušky tento problém obvykle řeší striktním nastavením kritérií hodnocení, která jsou rozdělena do různých oblastí, jimž je při hodnocení přisouzena různá bodová váha v závislosti na míře jejich funkčního naplnění.<sup>2</sup>

Tato práce se zaměřuje na výzkum záměrného analytického formativního hodnocení, resp. na měřitelné rozdíly mezi pedagogickým a vzájemným hodnocením. Obě formy hodnocení ve výzkumu rovněž vycházejí z předem nastavených kritérií.

---

<sup>1</sup> Jména zahraničních autorek nepřechyluji v textu ani v seznamu literatury.

<sup>2</sup> Příklad takového hodnocení představují např. písemné části státní maturitní zkoušky z češtiny nebo z cizího jazyka připravované Centrem pro zjišťování výsledků vzdělávání.

### 3 Formativní hodnocení

Stručný nástin rozdílu mezi sumativním a formativním hodnocením jsem uvedla v kapitole o typologii hodnocení. Tato kapitola se bude formativnímu hodnocení věnovat podrobněji. Uvedu v ní několik možných definic formativního hodnocení a popíšu některé etapy vývoje přemýšlení o formativním hodnocení. Blíže se zaměřím také na cíle, jejichž naplnění by tento typ hodnocení měl sloužit, a v neposlední řadě se budu věnovat také doporučením pro zavedení formativního hodnocení a shrnutí poznatků o této oblasti ve vztahu k výuce cizího jazyka.

#### 3.1 Vymezení pojmu

Pokud jde o vymezení pojmu formativní hodnocení, k dispozici se nabízí hned několik možností, z nichž asi nejširší nabízí dokument *Formative Assessment: Improving Learning in Secondary Classrooms* (OECD, 2005), který tento typ hodnocení charakterizuje následovně (tamtéž, s. 1):

„[Formativní hodnocení] identifikuje studijní potřeby žáků a reaguje na ně. Ve třídě, v níž se využívá formativní hodnocení, učitelé provádějí časté interaktivní hodnocení toho, jak studenti rozumějí látce. To jim umožňuje přizpůsobit jejich výuku tak, aby se setkala s individuálními potřebami studentů a lépe jim pomohla dosáhnout vysokých standardů. Učitelé studenty také aktivně zapojují do procesu a pomáhají jim rozvíjet dovednosti, které jim umožní lépe se učit“ (překlad AS).

Tato práce bude pracovat s poněkud užší definicí navrženou Blackem a Wiliamem (2009, s. 6), která pojem jednak vymezuje přesněji prostřednictvím škály, jednak klade důraz na možnosti různého zapojení jednotlivých aktérů:

„Hodnocení je formativní do té míry, do jaké jsou rozhodnutí o dalších krocích v procesu učení, která provádějí učitelé, studenti, nebo spolužáci na základě dokladů o výsledcích studentů elicitovaných a interpretovaných učiteli, studenty nebo spolužáky, lepší nebo lépe podložená oproti rozhodnutím, která se o takové podklady neopírají“ (překlad AS).

#### 3.2 Vývoj formativního hodnocení

Publikace, které se zabývají formativním hodnocením, můžeme pozorovat od konce 60. let 20. století, přičemž poměrně vysoký zájem o toto téma lze sledovat od 90. let.<sup>3</sup>

---

<sup>3</sup> Výběr studií v angličtině a španělštině převážně z období mezi lety 1998 a 2010 nabízejí např. Floréz a Sammons (2013), jejichž práce mapuje rozvoj formativního hodnocení v různých vzdělávacích kontextech na celém světě skrz celkem 33 odborných textů, z nichž 27 vzniklo právě v daném období (tamtéž, s. 5).

Práce zabývající se formativním hodnocením prošly mnohaletým vývojem od koncepce „mastery learning“, která byla ve středu zájmu v 70. a 80. letech minulého století, až po dnes převládající koncepci „learning for assessment“ (Torrance, 2012, s. 324). Oba tyto přístupy budu blíže charakterizovat v následujícím oddílu.

### **3.2.1 Mastery learning**

Formativní a sumativní hodnocení začalo být rozlišováno ve Spojených státech od 2. poloviny 60. let minulého století v rámci modelu tzv. „learning for mastery“, později „mastery learning“ (např. Bloom, 1968, 1971), do češtiny překládaného jako „zvládající učení“ (Pasch a kol., 1998, s. 112) nebo „učení směřující k zvládnutí“ (Průcha, 1997, s. 132).

Koncepce „mastery learning“ je založena na přesvědčení, že všichni žáci disponují schopností splnit stanovené cíle a že k tomu jen mohou potřebovat rozdílné množství času (Pasch a kol., 1998, s. 112). Jde o individualizovaný přístup k výuce, jehož klíčovým prvkem je stanovení dílčích cílů, které obsahují „všechny potřebné prvky (označení adresáta, chování podmínky a kritéria hodnocení)“ (tamtéž, s. 112).

Vedle těchto cílů je pro implementaci tohoto přístupu důležité také pravidelné využívání formativního hodnocení. To umožňuje jednak poskytovat studentům individualizovanou zpětnou vazbu a korektivní informace, které jim pomohou se zlepšit, jednak monitorovat postup plnění cílů a případně upravit výuku tak, aby lépe odpovídala potřebám studentů.

Vhodné se zdá být také využití pretestů, které učitelé poskytují předběžnou diagnostiku studentů – tedy vstupní informaci o tom, jaká je vzdálenost mezi počáteční znalostí studentů a stanovenými cíli. Důležitou částí Bloomovy teorie je totiž také jeho přesvědčení o tom, že ve výsledcích studentů hrají klíčovou roli kognitivní vstupní determinanty, tedy určité kognitivní vybavení (předpoklady pro učení), které student získal během předchozího učení (Průcha 1997, s. 134–135). V tomto ohledu tedy můžeme Bloomovu teorii „mastery learning“ zařadit do oblasti konstruktivistického vzdělání, jak ho popsal např. Tonucci (1991).

### **3.2.2 Assessment for learning**

„Assessment for learning“ představuje současný pohled na formativní hodnocení (Torrance, 2012, s. 324), jehož kořeny lze hledat ve článku *Inside the Black Box – Raising Standards Through Classroom Assessment* (Black – William, 2010 [1998]), v němž jeho autoři vyjadřují názor, že hodnocení představuje klíčový prvek výukového procesu a že je potřeba se proto intenzivně zabývat jeho kvalitou. Jejich příspěvek přitáhl k formativnímu hodnocení pozornost, a přestože se později objevila i kritika založená na jejich práci s daty (např. Benett, 2011),

navázaly na něj brzy další studie (např. Broadfoot a kol., 2002, Black a kol., 2003 nebo Black a kol., 2006).

„Assessment for learning“ můžeme přeložit jako „hodnocení pro učení“ nebo „hodnocení k podpoře učení“ (oba návrhy překladu AS). Jeho definici poskytují např. Broadfoot a kol. (2002, s. 2–3), kteří ho charakterizují jako „proces hledání a interpretace podkladů, které studentům a učitelům umožní rozhodnout se o tom, kde se v procesu učení nacházejí, jakým směrem potřebují jít a jak se tam nejlépe dostat“ (překlad AS). Vedle toho také uvádějí 10 principů, které by při jeho implementaci měly být dodrženy (tamtéž, s. 2–3). Blíže je představím v oddílu 3.3.

Termín „assessment for learning“ používají někteří autoři jako synonymum formativního hodnocení, jiní mezi těmito dvěma pojmy rozlišují (Wiliam, 2010, s. 24). Např. Black a kol. (2004, s. 10; citováno podle Wiliama 2010, s. 24) vidí rozdíl mezi formativním hodnocením a hodnocením k podpoře učení v tom, že hodnocení pro učení slouží žákovi k tomu, aby dosáhl lepších výsledků, zatímco formativní hodnocení slouží učitelům jako podklad pro rozhodnutí o tom, jak by měl případně upravit proces výuky. Tato práce se sice soustředí na užitečnost formativního hodnocení pro studenty, zároveň se ale domnívám, že hodnocení může vždy plnit i druhý zmíněný úkol – tedy poskytovat pedagogům informace potřebné k rozhodnutí o tom, jak bude výukový proces dál pokračovat. Zmíněné termíny budu proto v práci používat jako synonyma.

### **3.3 Zásady pro implementaci**

Stručný přehled praktických doporučení, která by se měla dodržovat v případě implementace formativního hodnocení, nabízí Broadfoot a kol. (2002, s. 2–3). Jejich přehled deseti principů (tamtéž, s. 2–3) blíže představím v následujících pododdílech.

V případě, že se pokusíme implementovat do hodin také vzájemné hodnocení nebo sebehodnocení, je podle mého názoru klíčové naučit tyto principy aktivně využívat i studenty.

#### **3.3.1 Efektivní plánování**

Učitel by měl dbát na to, aby studenti pravidelně získávali informace o svém pokroku a mohli je vztahovat směrem ke stanoveným učebním cílům.

Plánování by mělo zahrnovat strategie, které ověří, jestli studenti dobře rozumějí cílům výuky a chápou kritéria, podle nichž se bude jejich práce hodnotit. Mělo by se soustředit na to, jak studentům poskytne zpětnou vazbu, jak je do hodnocení zapojí a jak jim pomůže se dále zlepšovat.

### **3.3.2 Zaměření na proces učení**

Proces učení by měl být v centru pozornosti studentů i učitelů. Studenti by si měli být vědomi nejen vyučovacích obsahů, ale i způsobu, jakým si je osvojují.

### **3.3.3 Centrální postavení ve výukové praxi**

Charakteristiky hodnocení splňuje mnoho aktivit, které ve třídě probíhají. Učitelé a studenti by si tyto procesy měli uvědomovat a nezapomínat na to, že s nimi musejí pracovat všichni aktéři.

### **3.3.4 Klíčová dovednost učitele**

Formativní hodnocení by mělo být považováno za klíčovou dovednost profesionálního učitele. Učitelé by měli být podporováni v tom, aby rozvíjeli své dovednosti v oblasti plánování hodnocení, formulování zpětné vazby a podporování studentů v jejich sebehodnocení.

### **3.3.5 Citlivost a konstruktivnost**

Hodnocení může mít často na studenty emocionální vliv. Učitel by proto měl věnovat pozornost tomu, že komentáře a známky mohou ovlivnit sebedůvěru studentů, a být proto vždy empatický a poskytovat vždy co nejkonstruktivnější zpětnou vazbu.

Pomoci přitom může např. striktní používání popisného jazyka, který je vždy směřován k hodnocené práci, a nikoli k osobě studenta.

Domnívám se, že pokud chceme pracovat se vzájemným hodnocením, je potřeba, aby popisnou zpětnou vazbu nejprve studenti zvládli formulovat, např. na vzorových textech, a neohrozili tak vztahy ve skupině.

### **3.3.6 Podpora motivace**

Hodnocení by mělo studenty ovlivňovat příznivě. Učitel by proto vždy měl vědomě zdůrazňovat, že student dosáhl nějakého úspěchu nebo alespoň určitého pokroku, namísto toho, aby se soustředil pouze na chyby. Stejnou schopnost bychom v rámci formativního hodnocení měli budovat i u studentů, a rozvíjet tak jednak jejich schopnost dávat zpětnou vazbu jiným lidem, jednak přiměřeně hodnotit vlastní výkony a nepropadat negativismu.

Důležité je také poskytovat zpětnou vazbu na základě kritérií individualizovaně. Motivaci studentů totiž obvykle neprospívá, když učitel jejich výkony porovnává s výkony spolužáků, kteří dosáhli lepších výsledků. To může vést i k tomu, že se student přestane snažit v oblastech výuky, o nichž získal přesvědčení, že v nich není dost dobrý.

Učitel by v hodnocení navíc měl dopřát studentům určitou autonomii a dát jim příležitost k tomu, aby si alespoň částečně sami řídili svůj proces učení.



### **3.3.7 Porozumění cílům a kritériím**

Aby bylo využívání formativního hodnocení efektivní, je nezbytné, aby studenti předem porozuměli cílům a kritériím, podle nichž je hodnocena jejich práce. Na přijetí těchto cílů a kritérií může mít velmi pozitivní vliv, když se na jejich konstruování mohou studenti podílet.

### **3.3.8 Napomáhání seberozvoji**

Studenti potřebují vedení a informace, které jim pomohou pochopit, jak na sobě mohou pracovat, aby se dále zlepšovali. Učitel by jim proto během procesu učení měl poskytovat dostatek rad, které je povedou k tomu, aby si uvědomili své silné stránky a aby věděli o způsobech, jak pracovat na svých slabinách.

### **3.3.9 Rozvíjení dovednosti sebehodnocení**

Velkým přínosem správně implementovaného formativního hodnocení může být rozvoj metakognitivních dovedností studentů. Studenti se postupně mohou naučit lepší sebereflexi postupně převzít kontrolu nad svým učením.

### **3.3.10 Rozpoznávání všech výsledků vzdělání**

Učitel by si měl všímat nejen konečných výkonů, ale i celkových individuálních snah studentů a umožnit jim plně se zapojit do všech oblastí výuky. Měl by všem studentům umožnit co nejlépe naplnit jejich potenciál.

## **3.4 Formativní hodnocení a výuka cizích jazyků**

Formativní hodnocení ve výuce jazyka představuje nedílnou součást procesu učení. V jazykových školách na rozdíl od základních, středních a vysokých škol často nebývá kurz zakončen žádným testem a studenti v kurzech často nedostávají žádné výsledné hodnocení jako vysvědčení atp. Proto se domnívám, že by v jazykových kurzech mělo formativní hodnocení z principu převládat nad hodnocením sumativním, s tím, že lze pochopitelně očekávat určité rozdíly v kvalitě jeho implementace, např. v tom, do jaké míry jsou studenti předem seznámeni s kritérii, o jejichž splnění by měli usilovat, do jaké míry jsou vedeni k sebehodnocení, rozhodovací autonomii atd.

### **3.4.1 Výzkumy formativního hodnocení v jazykové výuce**

Formativnímu hodnocení je (nejen) v jazykové výuce poměrně velká pozornost věnována zejména v zemích, kde je postupné zavádění formativního hodnocení spojeno s reformami vzdělávacího systému v rámci vzdělávací politiky – mezi tyto země patří např. Austrálie, Kanada, Itálie, Anglie, Nový Zéland, Finsko, Dánsko (OECD, 2005, s. 1), Chile, Irán nebo

Hong Kong (Floréz – Sammons, 2013, s. 12). Metaanalýza, kterou provedly Floréz a Sammons, ukazuje, že se výzkumy obvykle zaměřují na tyto čtyři oblasti (tamtéž, s. 7–11): (i) dotazování (tj. jaké otázky učitel klade, jak důležitou roli mají tyto otázky ve výuce, jestli učitelé formulují otázky vhodně ve vztahu k jejich pochopitelnosti ze strany studentů, jestli učitel studentům poskytuje dostatek času na odpověď atd.), (ii) zpětná vazba a její důležitost (tj. např. nakolik informativní a popisná je zpětná vazba, nakolik je efekt takové zpětné vazby lepší než běžné, např. číselné známkování, nebo nakolik je důležité, aby zpětná vazba vycházela z předem daných kritérií), (iii) vzájemné hodnocení a sebehodnocení (tj. např. jaké tyto formy přinášejí studentům výhody v oblasti rozvíjení metakognice nebo učební autonomie nebo nakolik je jejich zavádění do výuky pro učitele náročné) a (iv) formativní využití sumativního hodnocení (tj. jestli jde o dvě různé kategorie hodnocení, nebo jestli jsou výsledky sumativního hodnocení využitelné také ve formativním procesu).

Výzkumy formativního hodnocení ve výuce cizího jazyka do velké míry kopírují oblasti popsané v předchozím paragrafu (srov. Chang, 2016). Častými dílčími otázkami může být např., jak studenti využívání této aktivity vnímají, nakolik důležitá je práce s kritérii a nácvik atd. Výsledky výzkumů zabývajících se vzájemným hodnocením, které jsou pro tuto práci nejvíce relevantní, popíšu detailněji v oddílu 4.1.

### **3.4.2 Hodnocení ve výuce češtiny jako cizího jazyka**

#### **Formativní hodnocení**

Pokud jde o didaktiku češtiny jako cizího jazyka, hodnocení se věnuje velká pozornost v souvislosti se snahami o systematickou implementaci SERR (2006). Ve vztahu k formativnímu hodnocení může být zajímavá např. práce na sebehodnotících deskriptorech pro češtinu jako cizí jazyk (Cvejnová – Šormová, 2018; Šormová, 2018). Tyto deskriptory lze při formativním hodnocení využít např. jako popisný nástroj sebehodnocení studentů nebo na jejich základě formulovat dílčí cíle výuky (Šormová, 2018, s. 175).

Formativnímu hodnocení ve výuce češtiny jako cizího jazyka se věnovala také Laufková (2018) ve svém příspěvku ve *Sborníku asociace učitelů češtiny jako cizího jazyka*. Její text ovšem nepopisuje formativní hodnocení specificky ve výuce češtiny, ale spíše zdůrazňuje užitečnost formativního hodnocení obecně a na příkladech ukazuje možnosti jeho implementace. Domnívám se, že příspěvek aspiruje na to, aby motivoval lektory k tomu, aby se o hodnocení více zajímali, ale že argumenty a rady v něm by byly pravděpodobně univerzálně platné i ve školním prostředí, a to i v jiných než jazykových výukových předmětech.

## Práce s chybou

Ve výuce češtiny jako cizího jazyka byla doposud poměrně velká pozornost věnována také samotné práci s chybou (srov. Švrčinová – Houžvička, 2014). Základní práci v tomto ohledu nejspíš reprezentuje Hrdlička (2012), který ve svém článku jednak pojednává o tom, co vlastně ve výuce češtiny ve vztahu k problematice jazykové správnosti, na niž v našem prostředí panují rozdílné názory, můžeme považovat za chybu (tamtéž, s. 89–101), jednak shrnuje dosavadní přístupy k typologii chyb a k jejich interpretaci (tamtéž, s. 101–104). Uvádí zde tyto druhy chyb: (i) „error“, tj. systematická chyba, které se student dopouští opakovaně a která je způsobená tím, že je daný jev nedostatečně internalizovaný (nový, nedostatečně pochopený...); (ii) „mistake“, tj. nestálá, kolísavá nebo občasná odchylka svědčící o nedůsledném učení se; (iii) „lapse“, tj. chyba, které se dopouštějí i rodilí mluvčí, pokud se dostatečně nesoustředí, jsou unavení nebo ve stresu; (iv) „careless slip“, tzv. přerěknutí, tj. chyba zapříčiněná nedbalostí (nesoustředěností, nedostatečným zájmem atd.).

Poslední část textu Hrdlička (tamtéž, s. 106–108) věnuje klasifikaci morfologických tvarů na (i) systémové, (ii) nesystémové a (iii) defektní a nabízí lektorům rady týkající se toho, jaké systémové tvary prezentovat studentům na vyšších a nižších úrovních zvládnutí jazyka. Hrdlička (tamtéž, s. 106–108) má v tomto ohledu za to, že studentům na nižších úrovních by se měly prezentovat převážně systémové tvary spisovné češtiny, zatímco pokročilejší studenti by měli být schopni nesystémové tvary identifikovat a adekvátně je používat v závislosti na komunikační situaci, v níž se pohybují.

## **4 Vzájemné hodnocení**

Základní rozdíl mezi vzájemným hodnocením, pedagogickým hodnocením a sebehodnocením jsem nastínila v kapitole 2.2 Typologie hodnocení. Tento oddíl se proto bude soustředit především na popis studií, které byly v této oblasti podniknuty, a na důsledky jejich zjištění pro výukovou praxi.

### **4.1 Dosavadní výzkumy vzájemného hodnocení**

Dosavadní výzkumy vzájemného hodnocení se soustředily jednak na spolehlivost tohoto typu hodnocení v porovnání se sebehodnocením nebo pedagogickým hodnocením (např. Langan a kol., 2008), jednak na reflexi těchto forem hodnocení ze strany studentů (např. Hanrahan – Isaacs, 2001 nebo Kaufmann – Schun, 2011). V oblasti jazykové výuky byl poměrně velký prostor věnován také vhodnosti použití této metody pro různé skupiny studentů a formulování zásad pro její úspěšné uplatnění (např. Villamil – de Guerrero, 1998). Výsledky některých studií detailněji popisují následující pododdíly.

#### **4.1.1 Známkování**

Poměrně velká pozornost byla ve výzkumu vzájemného hodnocení věnována jeho srovnávání s pedagogickým hodnocením a sebehodnocením co do spolehlivosti známek.

Redukovat debatu o vhodnosti použití vzájemného hodnocení pouze na to, jestli je srovnatelné s pedagogickým, nebere v úvahu jiné benefity, které studentům může toto nastavení přinést zejména v oblasti zlepšení jejich učení, která je v diskusi o formativním hodnocení klíčová (Liu – Carless, 2006, s. 279), a posouvá ji spíše do sféry sumativního hodnocení, které známky obvykle představují. Domnívám se ale, že výsledky studií, které se na srovnávání známek zaměřily, mohou do určité míry ukazovat tendence, na něž bychom měli při implementaci vzájemného hodnocení do výuky pamatovat.

Pokud jde o rozdíly ve známkování mezi pedagogickým a vzájemným hodnocením, metaanalýza několika textů, kterou zpracovaly Falchikov a Goldfinch (2000), podporuje tvrzení, že v průměru se známky, které dávají učitelé, se vzájemným hodnocením shodují (tamtéž, s. 315). Z jejich článku ale také vyplývá, že na shodu známek může mít vliv např. (i) povaha úkolu, (ii) míra seznámenosti studentů s kritérii hodnocení, (iii) jestli studenti známkují na základě kritérií globálně (větší shoda), nebo jestli známkují každé kritérium zvlášť (menší shoda), a samozřejmě i (iv) kvalita přípravy výzkumu, který známky porovnává (tamtéž, s. 315).

## **Pedagogické hodnocení, vzájemné hodnocení a sebehodnocení**

Co se týče sledovaných rozdílů ve známkování, z konkrétních studií zabývajících se srovnáním hodnocení na základě aktérů můžeme uvést např. výzkum Langana a kol. (2008). Tato studie porovnávala sebehodnocení, vzájemné hodnocení a pedagogické hodnocení studentů, kteří měli za úkol předvést krátkou prezentaci projektů, na nichž delší dobu pracovali (tamtéž, s. 180). Její výsledky ukázaly, že studenti měli tendenci dávat spolužákům vyšší hodnocení než pedagog, přičemž autoři tento jev vysvětlují blízkými vazbami v rámci daného kolektivu (tamtéž, s. 188).

Vedle pedagogického a vzájemného hodnocení bylo ve studii Langama a kol. (tamtéž) zkoumáno také sebehodnocení. Tato forma se od přechozích dvou výrazněji odlišovala. Během sebehodnocení měli totiž někteří studenti tendence své výkony oproti zpětným vazbám od pedagogů a kolegů z kurzu podceňovat. Autoři se domnívají, že na podhodnocování může mít určitý vliv gender a že k němu měly větší sklony studentky (tamtéž, s. 187).

Studie také ukázala, že při sebehodnocení studenti používali širší rozsah známek oproti vzájemnému hodnocení (tamtéž, s. 189). To může ukazovat také na sklon studentů nedělat mezi kolegy příliš velké rozdíly.

### **4.1.2 Vnímání vzájemného hodnocení ze strany studentů**

Další skupinu studií, které se zabývají vzájemným hodnocením, představují texty zkoumající postoje studentů k různým formám hodnocení. Mnoho z těchto textů ukazuje na to, že přestože má vzájemné hodnocení pozitivní vliv na rozvoj dovedností spojených s učením, mnoho studentů si vůči této formě výuky zachovává poněkud rezervovaný postoj, což souvisí s tím, že ho učitelé příliš často nevyužívají (Liu – Carless, 2006, s. 284). Jiné studie naopak ukazují, že přestože studenti v průběhu času spolu s přibývajícími zkušenostmi pocítují při vzájemném hodnocení větší sebedůvěru, zůstávají jim určité pochybnosti o jeho férovosti a konzistentnosti a mají obavy z vlivu přátelských vazeb nebo naopak ze soutěživosti plynoucí z nedostatečné anonymity (Smith a kol., 2002, s. 76). Mnoho těchto studií ovšem opět pracuje se studentským známkováním závěrečných prací a spadalo by proto opět nejspíš do kategorie sumativního hodnocení.

Detailněji se na problematiku přijetí vzájemného hodnocení dívají Kaufmann a Schun (2011), kteří v rámci rozsáhlejšího dotazníkového šetření zkoumali nejen to, jak studenti vnímají férovost známkování, ale i nápomocnost zpětné vazby, na jejímž základě mohou své práce přepracovat. Výsledky šetření ukazují, že přestože studenti skutečně projevují neochotu přijímat známky pouze na základě vzájemného hodnocení, považují komentáře spolužáků spíše

za přesné a spíše souhlasí s tím, že jim pomohly zlepšit se ve psaní (tamtéž, s. 397–398). Studie také naznačuje, že studenti vnímají vzájemné hodnocení jako užitečnější v kurzech, v nichž slouží jako doplněk pedagogického hodnocení, oproti kurzům, v nichž představuje jediný zdroj zpětné vazby (tamtéž, s. 404).

#### 4.1.3 Vzájemné hodnocení v jazykové výuce

O vzájemném hodnocení ve výuce češtiny se ve svém příspěvku zmiňuje Laufková (2018, s. 91), která jeho přínosy vidí (i) v jeho srozumitelnosti (spolužáci používají srozumitelný jazyk), (ii) v informační bohatosti (studenti dostanou zpětnou vazbu od několika spolužáků), (iii) v lepším přijetí (zpětná vazba od spolužáků může být příjemnější) a (iv) v pozitivním vlivu na motivaci studentů a zapojení celé třídy. Předpokladem pro jeho zavedení je podle ní ale bezpečné klima ve třídě a také dostatečný nácvik formulování zpětné vazby. Podle Laufkové (tamtéž) navíc není vhodné využívat tuto formu hodnocení ve všech typech aktivit, protože může vést k fixaci chyb.

Studie zaměřené na využití vzájemného hodnocení v jazykové výuce se z velké části věnují vysokoškolskému prostředí a zabývají se především výukou psaní (Cheng – Warren, 2005, s. 94). Metaanalýza, kterou podnikla Chang (2016) ukazuje, že se tyto výzkumy soustředí především na tyto tři dimenze (tamtéž, s. 86–106): (i) percepce, tzn. jaká přesvědčení a postoje studenti zaujímají ke vzájemnému hodnocení (obecně nebo třeba v závislosti na kultuře), (ii) proces, tzn. proces učení a implementace postupů vzájemného hodnocení, které zahrnují např. efekt nácviku, vztah mezi autorem a hodnotitelem, povahu zpětné vazby, komunikační jazyk, načasování pedagogické zpětné vazby nebo komunikační médium, (iii) produkt, tzn. k jakým výsledkům zahrnutí vzájemného hodnocení vedlo, pokud jde o výstupy výuky. Zjištění např. naznačují následující:

- (i) Studenti z asijských kultur, kde převládá kolektivismus<sup>4</sup>, často shledávají kritiku práce spolužáků jako více nepříjemnou (tamtéž, s. 88).
- (ii) Studenti, kteří mají možnost nácviku, dávají kolegům lepší zpětnou vazbu než studenti, kteří tuto možnost nemají a jejichž komentáře často trpí např. tím, že se zaměřují na dílčí problémy, nebo tím, že poskytují pouze vágní informace (tamtéž, s. 89–90).

---

<sup>4</sup> Rozdělení kultur na kolektivistické a individualistické vychází z typologie kultur Geera Hofstedeho (1983), v níž tato polarita představuje jednu ze čtyř dimenzí používaných pro obecné charakterizování kultur. Jde o dimenzi, která kulturu popisuje na základě míry těsnosti vztahů mezi lidmi. Zatímco v individualistických kulturách jsou vztahy mezi lidmi volnější a jednotlivci jsou orientováni na vlastní cíle, v kolektivistických kulturách jsou důležitější zájmy skupiny a vztahy mezi jednotlivými členy jsou proto v centru zájmu.

- (iii) Využití prvního jazyka studentů sice neumožňuje trénink užívání cílového jazyka, může ale napomoci např. v mluvené diskusi, která jinak může být ve smíšené třídě, v níž studenti nesdílejí první jazyk, náročná (tamtéž, s. 95–96).
- (iv) Učitel by měl zpětnou vazbu ke komentářům spolužáků zřejmě dávat až poté, co studenti jejich návrhy zahrnou (nebo naopak nezahrnou) do druhé verze svého textu, aby jejich rozhodnutí neovlivnil. Načasování zpětné vazby učitele by ale měla být ve výzkumu věnována větší pozornost (tamtéž, s. 96–97).
- (v) Zpětná vazba je užitečná jak v mluvené, tak v psané podobě. Zatímco mluvená zpětná vazba studentům umožňuje zaměřit se na globální výstavbu textu a na celkový obsah, psaná zpětná vazba jim umožňuje soustředit se detailněji na dílčí prvky textu a zjednodušuje jim zahrnutí návrhů úprav do druhé verze textu. Proto se někteří autoři domnívají, že je výhodné, aby studenti nejprve napsali písemnou zpětnou vazbu, kterou následně doplní o rozhovor nad ní (tamtéž, s. 99–100).
- (vi) Anonymizace zpětné vazby s využitím moderních technologií sice odbourává vliv sympatií a antipatií mezi studenty a vede k objektivnějšímu hodnocení, zároveň ale studentům může znemožnit pokládat ke komentářům doplňující otázky (tamtéž, s. 100–101).
- (vii) Míra, do jaké se studenti rozhodnou přijmout návrhy úprav od spolužáků, se napříč studii liší. Některé kvalitativní studie naznačují, že se studenti rozhodují (ne)akceptovat návrhy úprav na základě jejich kvality (neakceptují je např. v případech, kdy jsou komentáře vágní) nebo na základě toho, jestli s nimi souhlasí (tamtéž, s. 103–104).
- (viii) Pokud jde o zlepšení v psaní na základě vzájemného vs. pedagogického hodnocení, lze nalézt studie s protichůdnými výsledky. Této oblasti by se proto mělo ve výzkumu věnovat více pozornosti (tamtéž, s. 104).
- (ix) Některé výzkumy potvrdily pozitivnější dopady vzájemného hodnocení na hodnotitele oproti hodnocenému. Výzkumy by se v budoucnu proto měly zabývat touto oblastí více a pokusit se tento fenomén případně vysvětlit (tamtéž, s. 105).

## **4.2 Zásady pro implementaci vzájemného hodnocení**

Některé výzkumy, které do své metaanalýzy zahrnuly Floréz a Sammons (2013), zdůrazňují náročnost vzájemného hodnocení. Mnoho studentů podle nich není zvyklých přemýšlet o své práci ve vztahu k určitým výukovým cílům a je potřeba jim dát čas na to, aby si uvědomili, co

obnáší proces jejich učení, a dokázali na základě toho jednat. Trénink v hodinách proto musí zahrnovat definování kritérií a jejich jasnou komunikaci směrem ke studentům (tamtéž, s. 10).

Domnívám se, že zásady pro implementaci vzájemného hodnocení tedy kopírují principy, jež jsem popsala v oddílu 3.3, ale že je klíčový také nácvik jejich aktivního používání na straně studentů. Na význam dlouhodobějšího nácviku, během něhož studenti získávají potřebné zkušenosti, ukazuje např. metaanalýza několika studií, kterou provedli van Zundert a kol. (2010, s. 274–275).

#### **4.2.1 Jazyková úroveň**

Ve vztahu k jazykové výuce je vedle zmíněných principů samozřejmě potřeba brát v úvahu také úroveň, na níž se nacházejí vyjadřovací schopnosti studentů, protože poskytování zpětné vazby předpokládá např. schopnost vyjádřit a zdůvodnit svůj názor, dávat rady atd. Tyto kompetence studenti ovládají od úrovně B1 (SERR, 2006, s. 24).

Z hlediska této práce může být v tomto ohledu zajímavá např. studie *Assessing the Impact of Peer Revision on L2 Writing* (Villamil – de Guerrero, 1998), v níž autoři pracovali se středně pokročilými<sup>5</sup> vysokoškolskými studenty a sledovali, nakolik zpětnou vazbu od spolužáků inkorporují do druhých verzí svých textů. Autoři mají za to, že míra implementace návrhů spolužáků ukazuje na užitečnost vzájemného hodnocení ve výuce (tamtéž, s. 501). Ve vztahu k jazykové úrovni studentů je namístě uvést, že v rámci diskusí měli možnost využívat také svůj první jazyk (tamtéž, s. 500), a mohlo pro ně tedy být jednodušší komunikovat o textu a navrhovaných změnách tak, aby byla zpětná vazba srozumitelná.

#### **4.2.2 Sociální a věková skupina**

Jak už jsem naznačila, mnoho studií vzájemného hodnocení v jazykové výuce bylo provedeno ve vysokoškolském prostředí. Výjimkou v tomto ohledu může být např. studie *The Effect of Using Peer Assessment Training on Writing Performance among Arab EFL High School Students in Malaysia* (Almahasneh – Abdul-Hamid, 2019), která úspěšné využití této formy hodnocení pozorovala u studentů střední školy.

### **4.3 Online nástroje pro vzájemné hodnocení**

Vzájemné hodnocení lze ve třídě realizovat také s využitím různých online nástrojů, které nabízejí např. možnost přidělit studentům náhodně texty k ohodnocení a celý proces anonymizovat a zjednodušit jeho administrativu. Příkladem mohou být např. platformy jako

---

<sup>5</sup> Středně pokročilý (intermediate) obvykle referuje přibližně k úrovni B1.



Peergrade.io, Hypothes.is nebo Crowdgrader.org.<sup>6</sup> U některých těchto aplikací je využívání jejich bezplatné verze časově omezeno, vzájemné hodnocení lze ale realizovat např. i komentováním a následným sdílením souborů.

---

<sup>6</sup> Oproti původnímu plánu jsem se rozhodla srovnání online nástrojů do práce nezahrnout, protože jejím cílem je zhodnotit využitelnost metody, nikoli prozkoumávat nástroje, které její implementaci nejvíce usnadní.

## 5 Příprava výzkumu

### 5.1 Cíle výzkumu a jeho základní charakteristika

Výzkumná část si klade za cíl (i) ověřit využitelnost vzájemného hodnocení při přípravě studentů k písemné části zkoušky na úrovni B1 a (ii) zformulovat doporučení pro lektory, kteří by zavádění této metody měli v úmyslu. Za tímto účelem jsem provedla experiment, jehož cílem bylo srovnání pedagogického a vzájemného formativního hodnocení v intenzivním online kurzu psaní. Výzkumné otázky a hypotézy experimentu jsou představeny v oddílech 5.2 a 5.3. Oddíl 5.4 se detailně věnuje východiskům experimentu souvisejícím s požadavky zkoušky na úrovni B1.

Během experimentu si studenti vyzkoušeli dva nejčastější slohové útvary, které se u zkoušek na úrovni B1 vyskytují, přičemž jeden z útvarů byl hodnocen pedagogem, jeden studenti hodnotili vzájemně. Klíčové bylo u každé z úloh dodržení doporučeného formativní fázování úkolu – tzn. (i) pretest, na základě kterého bylo možné předběžně stanovit, na jaké úrovni znalostí se studenti nacházejí, (ii) společné vyvození kritérií na základě vzorových textů, (iii) vlastní psaní textu, (iv) rozbor zpětné vazby a (v) psaní druhé verze textu a implementace informací ze zpětné vazby.

Metodologicky se experiment opírá o podrobné kvalitativní analýzy: o chybovou analýzu<sup>7</sup> prvních verzí napsaných textů a o obsahovou analýzu hodnotících komentářů včetně analýzy jejich implementace v druhé verzi textu. Souborná vyhodnocení těchto analýz jsou provedena pomocí základních kvantitativních charakteristik. Závěrečnou částí je následné dotazníkové šetření mezi účastníky experimentu. Zpracování dat a vyhodnocení výzkumu je věnována šestá kapitola. Průběh experimentu a jeho reflexi podrobněji popisují v oddílech 5.5 a 5.6.

### 5.2 Výzkumné otázky

Ve vztahu k výzkumným cílům jsem si formulovala několik výzkumných otázek, které lze obecně rozdělit do dvou kategorií: otázky zkoumající (i) objektivní efektivitu metody a (ii) subjektivní hodnocení přínosu ze strany studentů. Jednotlivé otázky blíže popisují níže.

---

<sup>7</sup> Pojem „chybová analýza“ v této práci referuje k tabulce obsahující seznam jevů, které jsem označila za jednoznačné chyby, doplněný o stručnou charakteristiku jednotlivých položek ve vztahu k rovině textu.

### **5.2.1 Otázky týkající se objektivní efektivity metody**

#### **Jaké jsou porovnatelné odlišnosti v práci žáků s pedagogickým vs. vzájemným hodnocením?**

V rámci této otázky zkoumám, jak studenti pracují s komentáři, které obdrží od pedagoga, oproti komentářům, které dostanou od svých spolužáků. Zaměřuji se na to, kolik podnětů dostanou, a na to, kolik úprav na jejich základě provedou a v jakém rozsahu.

#### **Jak kvalitní hodnocení dokážou studenti na dané úrovni poskytnout?**

Tato otázka se zaměřuje na kompetence respondentů poskytovat kvalitní zpětnou vazbu. Vedle práce s komentáři se proto v analýze dívám i na to, jakých chyb si studenti všímají v textech svých kolegů a jaké naopak opomíjejí. Analyzuji také míru chyb v jejich úsudku, tzn.:

- (i) Identifikují chyby správně?
- (ii) Nabízejí správnou opravu?

### **5.2.2 Otázky týkající se subjektivního hodnocení přínosu ze strany studentů**

#### **Jak studenti hodnotí užitečnost a srozumitelnost obdrženého pedagogického a vzájemného hodnocení?**

V analýze dat se v rámci této otázky zaměřuji zejména na to, jak studenti subjektivně hodnotí oba typy hodnocení ve vzájemném srovnání, tzn. např. jestli spatřují nějaké výhody ve větším množství „recenzí“, které obdrželi od kolegů, jaké vnímají nevýhody ohledně jejich spolehlivosti atd. Sleduji také reflexi jejich zkušenosti v roli hodnotitele, tzn. např. jak studenti hodnotí možnost vidět/komentovat texty svých kolegů nebo jak studenti hodnotí náročnost psaní komentářů k textům svých kolegů.

## **5.3 Výzkumné hypotézy**

Ve vztahu k již popsaným výzkumným otázkám jsem na základě nastudované literatury, kterou jsem podrobněji rozebrala v teoretické části, a podle své dosavadní zkušenosti z výuky zformulovala také hypotézy, které pro větší přehlednost uvádím spolu s otázkami, k nimž se vztahují.

#### **Jaké jsou porovnatelné odlišnosti v práci studentů s pedagogickým vs. vzájemným hodnocením?**

- 1) Respondenti dostanou více podnětů od lektora než od svých spolužáků.
- 2) Respondenti provedou méně úprav na základě komentářů od kolegů než na základě pedagogického hodnocení.

### **Jak kvalitní hodnocení dokážou studenti na dané úrovni poskytnout?**

- 1) Studenti neurčí správně všechny nedostatky v textech svých kolegů.
- 2) Studenti nebudou ke všem chybám vždy nabízet správnou opravu.

### **Jak studenti hodnotí užitečnost a srozumitelnost obdrženého pedagogického a vzájemného hodnocení?**

- 1) Studenti budou hodnotit pedagogické hodnocení jako užitečnější.
- 2) Studenti budou pozitivně hodnotit možnost vidět texty spolužáků.
- 3) Studenti budou pozitivně hodnotit větší množství recenzí, které obdrží v rámci vzájemného hodnocení.

## **5.4 Východiska experimentu**

V této části kapitoly nejprve stručně charakterizují úroveň B1 a popíšu strukturu dvou jazykových zkoušek, které studenti na této úrovni mohou skládat. Pozornost přitom věnuji především oblasti psaní, která je pro tento výzkum klíčová. V dalším oddílu pak shrnu dostupné informace o kritériích hodnocení písemných prací u zmíněných zkoušek.

### **5.4.1 Úroveň B1**

#### **Globální popis**

*Společný evropský referenční rámec pro jazyky* (SERR, 2006, s. 23–24, 34) popisuje úroveň B1 jako tzv. „prahovou úroveň“, která představuje přechod od jednoduchého používání jazyka v situacích, které se mluvčího týkají, směrem k pružnému zvládnutí zcela samostatné každodenní komunikace. I proto je zřejmě tato úroveň požadována např. pro vstup do některých zaměstnání (Boccou Kestřánková a kol., 2013, s. 8) nebo pro udělení státního občanství.

SERR (2006, s. 24) dále uvádí, že student na této úrovni (i) „[r]ozumí hlavním myšlenkám srozumitelné spisovné vstupní informace týkající se běžných témat, se kterými se pravidelně setkává v práci, ve škole, ve volném čase atd.“, (ii) „[si umí] poradit s většinou situací, jež mohou nastat při cestování v oblasti, kde se tímto jazykem mluví“, (iii) „[u]mí napsat jednoduchý souvislý text na témata, která dobře zná nebo která ho/ji osobně zajímají“ a (iv) „[d]okáže popsat [události, své zážitky], sny, naděje a cíle a umí stručně vysvětlit a odůvodnit své názory a plány.“

Témata, o nichž se zmiňuje bod (i), blíže popisuje např. publikace *Prahová úroveň – Čeština jako cizí jazyk* (Bischofová a kol., 2001, s. 155–209), která k nim nabízí i konkrétnější deskripci a okruhy slovní zásoby.

Stručný seznam témat uvádím zde (tamtéž, s. 155–209):

- (i) osobní údaje,
- (ii) dům, domácnost, nejbližší okolí,
- (iii) každodenní život,
- (iv) volný čas, zábava,
- (v) cestování,
- (vi) vztahy s lidmi,
- (vii) zdraví,
- (viii) vzdělání,
- (ix) nákupy,
- (x) jídlo a pití,
- (xi) služby,
- (xii) hledání cesty a místa,
- (xiii) počasí

### **Písemný projev**

Pokud jde o bližší pohled na psaný projev studentů, charakterizuje ho SERR (2006, s. 26) následovně: Uživatel jazyka umí „napsat jednoduché souvislé texty na témata, která dobře [zná] nebo která [ho] osobně zajímají. [Umí] psát osobní dopisy popisující zážitky a dojmy.“

Tento popis dále rozvádějí Bischofová a kol. (2001, s. 217–218), podle nichž jsou možnosti studentů na této úrovni stále ještě relativně omezené. Studenti podle nich zatím „nepřekročí schopnost sestavit jednoduchá písemná sdělení“, která následovně specifikují zejména jako (i) jednoduché formuláře potřebné např. při cestování, (ii) „standardní dopisy s účelovým zaměřením“ týkající se např. žádosti o informace a (iii) „jednoduchou osobní korespondenci (od psaní krátkých pozdravů přes jednoduchá sdělení až k jednoduchým osobním dopisům).“ Uživatelé jazyka na této úrovni by proto měli vědět: (i) jak zvládnout výstavbu a uspořádání dopisu, (ii) jak vhodně používat úvodní a závěrečné formule, které se v dopisech využívají, (iii) jak správně napsat datum a (iv) jak správně používat malá a velká písmena a také interpunkční znaménka.

### **5.4.2 Zkouška z češtiny pro udělení státního občanství**

#### **Právní úprava**

*Zákon č. 186/2013 Sb., o státním občanství České republiky* (§ 14) stanovuje uchazečům o české státní občanství povinnost prokázat znalost českého jazyka. To se netýká pouze

žadatelů, kteří i) doloží, že absolvovali alespoň 3 roky základní, střední nebo vysoké školy, kde byla vyučovacím jazykem čeština, ii) nedosáhnou ke dni podání žádosti 15 let nebo naopak dosáhnou více než 65 let nebo iii) mají tělesný nebo mentální handicap, který jim neumožňuje si znalost českého jazyka osvojit.

Podobu zkoušky z českého jazyka pak dále upravuje *Vyhláška č. 433/2013 Sb., o prokazování znalosti českého jazyka a českých reálií pro účely udělování státního občanství České republiky*, která stanovuje úroveň zkoušky jako úroveň B1 podle SERR (§ 2), hranici úspěšnosti (§ 8) a blíže specifikuje také další otázky, zejména organizačního charakteru, jako je např. proces přihlašování (§ 5), uchovávání protokolů (§ 7) nebo seznam rovnocenných zkoušek, které lze případně uchazečům o občanství uznat jako ekvivalentní (§ 13).

### **Celková podoba zkoušky**

Zkoušku z českého jazyka pro udělení českého státního občanství připravuje a organizuje Ústav jazykové a odborné přípravy Univerzity Karlovy (dále ÚJOP UK).

Podle informací dostupných na oficiálních webových stránkách Národního pedagogického institutu České republiky (Zkouška pro občanství, 2020; Zkušební řád, 2021, s. 5–6) se zkouška skládá z písemné a ústní části. Písemná část sestává ze tří subtestů, v nichž mají uchazeči prokázat zvládnutí čtení s porozuměním, poslechu s porozuměním a také písemného projevu na úrovni B1. Ústní část zkoušky je realizována prostřednictvím krátkého pohovoru.

Představu o přesnější podobě zkoušky lze získat z komentované modelové varianty zkoušky (Modelová varianta, 2019), která pro kandidáty představuje jediný zdarma dostupný vzorový test. Shrnutí podává tabulka 1.

**Tabulka 1: Modelová varianta zkoušky z češtiny pro udělení státního občanství**

Název části	Úloha	Rozsah textu	Počet úkolů	Forma úkolu	Počet bodů
<b>Čtení s porozuměním</b> 50 minut	Úloha 1	1 středně dlouhý text (100–110 slov)	6	Výběr ze dvou alternativ (ano/ne)	6
	Úloha 2	1 dlouhý text (300–330 slov)	8	Výběr ze čtyř alternativ	8
	Úloha 3	6 krátkých textů (celkem 250–275 slov)	6	Přiřazování	6
	Úloha 4	5 částí jednoho textu (celkem 150–165 slov)	5	Uspořádávání	5
<b>Poslech s porozuměním</b> 30–35 minut	Úloha 1	3–6 krátkých dialogů (celkem 250–275 slov)	6	Výběr ze čtyř alternativ	6
	Úloha 2	1 dlouhý dialog (350–385 slov)	8	Výběr ze čtyř alternativ	8
	Úloha 3	1 středně dlouhý monolog/dialog (200–220 slov)	6	Výběr ze dvou alternativ (ano/ne)	6
	Úloha 4	2–5 krátkých monologů (celkem 100–110 slov)	5	Otázky se stručnou odpovědí	5
<b>Psaní</b> 60 minut	Úloha 1	Produkováný text min. 75 slov	1	Otevřená úloha	12
	Úloha 2	Produkováný text min. 100 slov	1	Otevřená úloha	13
<b>Ústní část (mluvení)</b> 15–18 minut/dvojice	Úloha 1	2 minuty / kandidát		Řízený monolog	6
	Úloha 2	3 minuty / kandidát		Řízený konverzace s examínátorem na vybrané téma	6
	Úloha 3	5–8 minut / dvojice		Dialog kandidátů	6
	Ústní interakce			(hodnotí se celkem za všechny úlohy)	3
	Fonologická kompetence			(hodnotí se celkem za všechny úlohy)	4

Zdroj: Modelová varianta (2019, s. 6).

### Subtest psaní

Jak vyplývá z tabulky 1, subtest psaní se skládá ze dvou úloh, na jejichž vypracování mají kandidáti celkem 60 minut. V obou úlohách má kandidát prokázat, že „dovede napsat jednoduše členěné souvislé texty na známá témata, sdělit informace a myšlenky, popsat zážitky a události a vyjádřit své pocity a názory“ (tamtéž, s. 36).

V první úloze (tj. např. dotazník hodnotící služby, reklamující zboží atd.) musí uchazeč napsat text o minimálně 75 slovech, ve druhé úloze (tj. např. neformální dopis, zpráva,

vyprávění) text o 100 slovech (tamtéž, s. 36). Ve druhé úloze si student vybírá ze dvou nabízených témat (tamtéž, s. 22).

Ačkoli se komentář k modelové variantě zmiňuje o různých typech textu, z informací, které jsem dostala od pracovníků ÚJOP UK<sup>8</sup> a od kolegyň, které se přípravou ke zkoušce dlouhodobě zabývají,<sup>9</sup> vyplývá, že u zkoušky se opakovaně vyskytují především dotazníky v úloze č. 1 a dopisy, v nichž mají kandidáti vyjádřit svůj názor na nějaké téma, v úloze č. 2.

### 5.4.3 Certifikovaná zkouška z češtiny CCE-B1

Možným ekvivalentem zkoušky pro udělení státního občanství je pro uchazeče *Certifikovaná zkouška z češtiny na úrovni B1*, dále CCE-B1 (Informace o podobě zkoušky, 2019, s. 3). I tyto zkoušky připravuje a organizuje ÚJOP UK, a možná proto je jejich podoba téměř totožná. Jak vyplývá z aktualit na webových stránkách ÚJOP UK a z informací, které jsem získala od pracovníků ÚJOP,<sup>10</sup> zkouška častěji podléhá dílčím změnám, které nebývají oficiální cestou předem hlášeny. I u ní ale studenti během písemné a ústní části prokazují zvládnutí poslechu a čtení s porozuměním a produktivní dovednosti v písemném a mluveném ústním projevu.

#### Celková podoba zkoušky

Informace o organizaci zkoušky podává tabulka 2.

**Tabulka 2: Struktura certifikované zkoušky CCE-B1**

	Čtení s porozuměním		Poslech s porozuměním		Psaní		Mluvení
Počet úloh	4	<i>Přestávka</i>	4	<i>Přestávka</i>	2	<i>Přestávka</i>	3
Počet úkolů	25	<i>minimálně</i>	25	<i>minimálně</i>	-	<i>minimálně</i>	-
Délka zkoušky	50 minut	<i>10 minut</i>	35–40 minut	<i>10 minut</i>	60 minut	<i>20 minut*</i>	15–18 minut (2 kandidáti)
Maximální počet bodů	25		25		25		25
Hraniční skóre (60 %)	15		15		15		15

\* Pořadí kandidátů se losuje. Zdroj: Informace o podobě zkoušky (2019, s. 5).

Detailnější popis zkoušky CCE-B1 (tamtéž, s. 5–10) pak naznačuje, že forma testových úloh je do velké míry podobná a rozdíly jsou spíše dílčí a vzešly z úprav, které ÚJOP postupně provádí (jedná se např. o množství variant, z nichž kandidát vybírá správnou odpověď, o počet

<sup>8</sup> Osobní sdělení Kateřiny Vodičkové a Martiny Hulešové ze dne 1. 2. 2021.

<sup>9</sup> Např. osobní sdělení Lenky Krškové ze dne 12. 3. 2021.

<sup>10</sup> Osobní sdělení Kateřiny Vodičkové a Martiny Hulešové ze dne 1. 2. 2021.



úloh u jednotlivých textů atd.). Tyto rozdíly nejsou pro tuto práci příliš relevantní, a proto je nebudu rozebírat detailněji.

### **Subtest psaní**

Subtest psaní se u CCE-B1 sestává obdobně jako u zkoušky pro udělení státního občanství ze dvou úloh, na jejichž vypracování má kandidát celkem 60 minut.

Úlohy blíže charakterizuje *Manuál pro přípravu k subtestu psaní* (2020), který již reflektuje změny platné od roku 2021 a který obsahuje jedinou aktuální modelovou verzi zkoušky, ke které se uchazeči mohou dostat zdarma. První úlohu i tady představuje vyplnění dotazníku o třech otázkách, v odpovědi na něž mají studenti napsat celkem alespoň 75 slov. Druhou úlohou je opět neformální dopis kamarádovi nebo kamarádce o minimálně 100 slovech (tamtéž, s. 2). I zde si kandidát může vybrat ze dvou témat, ta se ale zřejmě nově poněkud odlišují od původní varianty – zatímco dříve měl kandidát vyjádřit svůj názor na nějaké obecné tvrzení (stejně jako u zkoušky pro získání českého občanství), v nových verzích zkoušky uchazeči popisují fiktivní situaci, do níž se dostali a kterou musí řešit. Jako příklad pro srovnání lze uvést dvě zadání z modelových testů:

- (i) „Jídlo je radost“ (Boccou Kestřánková a kol., 2013, s. 36).
- (ii) „Koupil/a jste si nové auto a máte s ním problém“ (*Manuál pro přípravu k subtestu psaní*, 2020, s. 15).

### **5.4.4 Kritéria hodnocení písemných prací**

#### **Kritéria hodnocení**

Kritéria hodnocení jsou pro zkoušku pro udělení státního občanství detailně popsána v okomentované *Modelové variantě* (2019, s. 6), pro zkoušku CCE-B1 v *Informacích o podobě zkoušky* (2019, s. 7–8). V případě zkoušky CCE-B1 jsou navíc v *Manuálu pro přípravu k subtestu psaní* (2020, s. 7–16) k dispozici okomentované ukázky textů, jež u zkoušky uspěly, nebo naopak neuspěly. Ze všech dostupných dokumentů lze vyvodit následující kritéria, která jsou společná pro hodnocení obou zkoušek:

- 1) míra splnění úkolu (dodržení tématu, typu textu a minimálního rozsahu)
- 2) míra zdařilosti výstavby textu (koherence a koheze)
- 3) míra rozsahu a správnosti použití slovní zásoby (vzhledem k popisu úrovně B1)
- 4) míra rozsahu a správnosti použití gramatických prostředků (vzhledem k popisu úrovně B1)

## **Bodování**

U obou zkoušek může kandidát za obě písemné úlohy získat v součtu maximálně 25 bodů: v úloze č. 1 maximálně 12 bodů, v úloze č. 2 maximálně 13 bodů. Hranice úspěšnosti je stanovena na 15 bodů, tj. 60 % (srov. Informace o podobě zkoušky, 2019, s. 5, 7; Modelová varianta, 2019, s. 6; Zkušební řád, 2021, s. 9).

ÚJOP UK na rozdíl např. od Centra zjišťování výsledků vzdělávání (CERMAT) nezveřejňuje ani neposkytuje k nahlédnutí konkrétní bodovací tabulku s kritérii, podle níž se písemné práce studentů hodnotí. Z osobního sdělení pracovníků ÚJOP UK vyplývá, že tým, který zkoušky připravuje a hodnotí, je přesvědčen o tom, že zveřejnění podobného dokumentu by bylo nežádoucí, protože bez příslušného školení neposkytuje lektorům jazykových škol oporu pro hodnocení cvičných textů jejich studentů. Jeho únik by navíc podle nich mohl zvýhodnit některé jazykové školy oproti jiným.<sup>11</sup>

K dispozici jsou tedy pouze dílčí informace. Text se například nehodnotí dál, pokud autor nedodrží téma, nedostatečně zpracuje zadání nebo napíše méně než 55 slov v úloze č. 1 nebo méně než 70 slov v úloze č. 2 (Informace o podobě zkoušky, 2019, s. 8). Chyby v gramatice a v lexiku jsou pak posuzovány jako různě závažné podle míry, do jaké znesnadňují porozumění (tamtéž, s. 7).

### **5.4.5 Reflexe zkoušek v podobě experimentu**

Obsah kurzu a kritéria hodnocení vycházely z podoby zkoušek, které jsem popsala v oddílu výše. Studenti se během experimentu podrobněji seznámili se zásadami pro psaní dopisu, v němž vyjádří svůj názor, a se zásadami pro psaní odpovědí do dotazníku spokojenosti. Kritéria hodnocení zahrnovala (i) dodržení instrukcí, (ii) logickou výstavbu argumentů a návaznost textu, (iii) rozsah a správnost gramatiky, (iv) rozsah a správnost lexika. Studenti měli na psaní textů také časový limit, avšak vzhledem k tomu, že se jednalo o výuku (nikoli o „zkoušku nanečisto“), a úlohy byly navíc vypracovávány zvlášť, byl tento limit pozměněn, resp. navýšen (podrobněji v oddílu 5.5).

## **5.5 Organizace a popis experimentu**

Experiment se uskutečnil v rámci intenzivního kurzu psaní pro uchazeče o zkoušku B1, který proběhl online na platformě Zoom<sup>12</sup> během dvou po sobě jdoucích víkendů, z nichž byl každý věnován jedné z písemných úloh, jež se pravidelně objevují jak u certifikované zkoušky

---

<sup>11</sup> Osobní sdělení Kateřiny Vodičkové a Martiny Hulešové ze dne 1. 2. 2021.

<sup>12</sup> K online kurzu bylo potřeba přistoupit vzhledem k platným nařízením vlády proti šíření covidu-19, která v březnu 2021 neumožňovala konání prezenčních jazykových kurzů.

z češtiny CCE-B1, tak u jazykové zkoušky pro udělení českého státního občanství, tzn. (i) neformální dopis během prvního víkendu a (ii) dotazník spokojenosti během druhého víkendu. V obou víkendech jsem práci se studenty rozdělila do pěti fází. Tyto fáze představuji v následujících pododdílech.

### 5.5.1 První fáze: Stanovení kritérií

První fáze práce se studenty vždy spočívala ve společném konstruování kritérií v řízené diskusi. Jako opěrný materiál pro ni sloužily příkladové texty, v nichž měli studenti hledat silné a slabé prvky. Příkladové texty se skládaly jednak z upravených textů od dalších mých studentů na úrovni B1, kteří souhlasili s jejich využitím ve výuce, jednak z upravených textů, které jsem převzala z nejnovější příručky pro přípravu k certifikované zkoušce CCE-B1. Texty tedy vzhledem k úpravám nepředstavují autentickou produkci, ale jsou adaptované tak, aby upoutaly pozornost studentů na jednotlivá kritéria. V některých textech jsem tak např. opravila gramatiku, a naopak zdůraznila nelogičnost argumentů, jiné texty jsem výrazně zjednodušila ve slovní zásobě a syntaktických konstrukcích atd. Jednotlivé texty uvádím v příloze č. 1.

Po rozboru textů následovala společná sumarizace toho, jak by měl vypadat ideální text, který by ve zkoušce uspěl, doplněná o mé shrnutí v bodech (příloha č. 1). Tuto sumarizaci jsem dále doplnila plakátem *10 tipů jak uspět v psaní aneb Pozor na nejčastější chyby* (příloha č. 2), který je volně přístupný na stránkách ÚJOP UK.

### 5.5.2 Druhá fáze: Psaní úlohy

Ve druhé fázi kurzu měli studenti během stanoveného časového limitu za úkol napsat podle zadání (příloha č. 1, str. 10, 24) text daného typu. V případě dopisu měli na výběr ze dvou témat, v případě dotazníku bylo stanoveno pouze jedno téma, a sice hodnocení jazykového kurzu.<sup>13</sup>

Časový limit jsem u úloh stanovila na 45 minut na dopis a na 25 minut pro dotazník. Během zkoušky mají studenti k dispozici 60 minut na obě úlohy, ale vzhledem k tomu, že jsme texty procvičovali odděleně, zdálo se mi vhodné čas rozdělit nerovnoměrně podle délky a náročnosti formy. Studenti dále dostali pět minut navíc u každé úlohy, protože nešlo o zkoušku nanečisto, ale o nácvik dovedností.

Úlohy studenti psali rukou na papír a neměli přitom dovoleno používat žádné pomůcky. Hotové texty odevzdávali jako fotografickou přílohu prostřednictvím emailu. Většina studentů obě úlohy dokončila pět až deset minut před koncem časového limitu, někteří studenti i dříve.

---

<sup>13</sup> Po kurzu měli studenti za úkol vyplnit další dotazník, v němž měli kurz zhodnotit. Zadání úlohy jim tedy následně mohlo pomoci s formulací zpětné vazby pro účely výzkumu.

Studenti, kteří využili celý časový limit, obvykle měli buďto technické problémy, nebo napsali výrazně delší text, než zadání požadovalo.

### 5.5.3 Třetí fáze: Hodnocení

Třetí fázi kurzu představovalo hodnocení napsaných textů. K dopisům jsem poskytovala zpětnou vazbu já, dotazníky si studenti hodnotili vzájemně.

#### Pedagogické hodnocení

Vzhledem k nízké kvalitě některých fotografií jsem rukou psané texty převedla do textové formy a následně je komentovala a navrhovala úpravy v režimu sledování změn v textovém procesoru MS Word. Texty jsem následně nahrála na sdílené úložiště Google Drive, abych se ujistila, že si studenti budou moci všechny mé komentáře a úpravy zobrazit i na svých zařízeních v nezměněné podobě.

Je na místě upřesnit, že ačkoli byla velká část textů na úrovni B1 a písemný projev respondentů ji mohl v některých aspektech přesahovat, přistoupila jsem k opravám všech případných nedostatků bez ohledu na jazykovou úroveň. Během reflexe zpětné vazby jsem na tento postup studenty ovšem upozornila, abych se vyhnula případné demotivaci.

Ohodnocené texty jsou uvedeny v příloze č. 4.

#### Vzájemné hodnocení

Pro vzájemné hodnocení jsem zvolila text, který je kratší a typově jednodušší, tj. dotazník. Do druhého víkendu jsem ho zařadila proto, aby studenti měli možnost se v kurzu trochu poznat a aby měli možnost na textech opravených mnou vidět, jak mohou postupovat. Samotnému vzájemnému hodnocení ale předcházela samozřejmě i další příprava. Během druhého víkendu jsem např. v první fázi práce studentům u jednotlivých textů vždy kladla otázku *Co můžeme autorovi doporučit, aby byl jeho text příště lepší?* a po této fázi jsem zařadila ještě další aktivitu, během níž jsme se studenty diskutovali o několika různých příkladech hodnocení a společně formulovali a následně shrnuli několik zásad pro psaní komentářů (příloha č. 1, str. 25–28).

Každý student hodnotil celkem 3 texty svých kolegů. Hodnocení probíhalo prostřednictvím sdílených dokumentů na Google Drive formou přidávání komentářů k PDF dokumentu.<sup>14</sup> Vzhledem k minimu času, který jsem měla k dispozici, jsem texty ponechala ve formě rukopisů, studenty jsem ale předem instruovala, aby posílali fotografie v lepší kvalitě.

---

<sup>14</sup> Využití sdílených dokumentů přes Google Drive jsem zvolila, protože mi umožnilo nechat studenty komentáře přidávat k ručně psaným textům, je pro studenty jednoduché, protože s ním mnohdy již mají zkušenosti, a schopnost studentů s ním pracovat lze během online výuky poměrně snadno otestovat v režimu sdílené obrazovky. Okomentované dokumenty je pak navíc možné jednoduše stáhnout tak, aby je měl lektor možnost dále analyzovat.

Komentáře i hodnocené texty byly anonymizované, abych eliminovala vliv případných sympatií a antipatií mezi studenty a zajistila větší pocit bezpečí. Přidělení textů ale nebylo náhodné. Na základě pretestů, dopisů a předchozí interakce jsem zhodnotila jazykovou úroveň jednotlivých studentů a pokusila jsem se texty k hodnocení přidělit tak, aby každý student obdržel alespoň jednu recenzi od spolužáka s vyšší úrovní jazykového projevu a zároveň aby měl každý student možnost vidět alespoň jeden text vyšší kvality.

Studenti měli na hodnocení textů spolužáků celkem 70 minut a většina z nich skončila zhruba 15 minut před koncem limitu. V případě jednoho z respondentů jsem ale musela časový limit prodloužit o 30 minut, protože v opačném případě by některé texty zůstaly zcela neokomentované.

Okomentované texty jsou uvedeny v příloze č. 5.

#### **5.5.4 Čtvrtá fáze: Diskuse nad obdrženým hodnocením**

Ve čtvrté fázi, která proběhla až následující den, studenti dostali k dispozici své texty doplněné o komentáře a návrhy úprav a mohli vznášet konkrétní otázky buďto soukromě v chatu, nebo nahlas před dalšími studenty.

V případě pedagogického hodnocení předcházela této diskusi ještě reflexe některých chyb. Nejčastější chyby v textech se týkaly gramatiky (zejména tvarosloví) nebo pravopisu (zejména kvantity). Vzhledem k tomu, že odstranění těchto nedostatků obvykle vyžaduje dlouhodobý nácvik, zaměřila jsem se v kurzu na některé méně frekventované, ale také opakované chyby (slovosled, opakování slov) a pokusila jsem se studentům vysvětlit principy a kompenzační strategie, jejichž znalost jim může pomoci je odstranit (podrobněji viz příloha 1, str. 12–13).

Vzájemné hodnocení bylo anonymní, a studenti tudíž následně neměli možnost klást kolegům doplňující dotazy k jednotlivým komentářům. Vzhledem k tomu, že kvalita komentářů byla velmi nevyrovnaná a některé z nich např. identifikovaly chyby nesprávně nebo navrhovaly nevhodné opravy, jsem ale studentům umožnila položit případné doplňující dotazy mně. V rámci chatu mi ovšem studenti položili jen sedm jednotlivých dotazů (z toho byly čtyři od jednoho respondenta), a na následující přepracované verze textu tedy měly moje odpovědi jen minimální vliv. V případě větších změn jsem vliv konzultace zaznamenala do poznámek.

#### **5.5.5 Pátá fáze: Druhá verze textu**

V poslední fázi měli studenti za úkol na základě obdrženého hodnocení přepracovat první verzi textu. To se zprvu potkalo u některých studentů s nepochopením, nakonec ale souhlasili, že to pro ně může být užitečné. V případě několika respondentů ale můžeme přesto vypožorovat, že

rozdíl mezi první a druhou verzí textu výrazně překročil doporučení, která student obdržel v rámci hodnocení (srov. příloha č. 4, str. 7–8, 10 a příloha č. 6, str. 3–4).

Studenti také dostali možnost napsat si doma ve svém volném čase stejný typ textu na jiné téma a poslat mi ho dodatečně mimo kurz e-mailem k ohodnocení. Tuto možnost ale nakonec v obou případech využili jen tři studenti.

## **5.6 Dotazník**

Po skončení kurzu měli studenti za úkol vyplnit dotazník, v němž měli subjektivně zhodnotit užitečnost kurzu a použitých metod.

První část dotazníku obsahovala tři obecné otázky, které kopírovaly obvyklou strukturu písemné úlohy u zkoušky. Studenti je obvykle využili ke zhodnocení formátu kurzu, atmosféry v kurzu, práce lektorky, složení skupiny a tak dále (detailněji viz příloha č. 8).

Další oddíly otázek mířily konkrétně na použité metody. Studenti v nich měli formulovat výhody a nevýhody pedagogického a vzájemného hodnocení. Na škále měli také hodnotit užitečnost a srozumitelnost komentářů. Podobným způsobem měli zhodnotit také to, jak na ně působila zkušenost v roli hodnotitele jiných textů (konkrétní znění otázek viz příloha č. 8).

## **6 Vyhodnocení výzkumu**

### **6.1 Respondenti**

Do víkendového kurzu psaní se přihlásilo celkem 11 zájemců, kteří potvrdili, že se v současné době připravují na zkoušku z češtiny pro udělení státního občanství, přičemž dva z nich v průběhu kurzu ze zdravotních a osobních důvodů ukončili účast v experimentu. Jejich data jsem proto z analýzy vyřadila.

Všichni respondenti před vstupem do kurzu podepsali informovaný souhlas s účastí ve výzkumu, konzultovaný s právníčkou. Texty získané v rámci experimentu byly následně anonymizovány, přičemž v případě dopisů jsem pristoupila i ke změně jmen použitých v podpisu, tak aby odpovídaly jménům běžně užívaným v zemi původu respondentů.

Vhodnost kurzu pro studenty jsem ověřila pomocí pretestů, jejich jazykovou úroveň jsem ovšem detailněji netestovala. Ze získaných dat ale vyplynulo, že přestože velká část studentů v některých aspektech téměř přesahuje úroveň B1, mnoho z nich vykazuje nedostatečné osvojení gramatiky a také pravopisných a hláskoslovných jevů. Celkové rozpětí se podle mého odhadu pohybovalo zhruba na škále A2-B2, přičemž většina studentů se pohybovala v jejím středu. Z mé zkušenosti se však tato situace výrazně neliší od běžné praxe v přípravných jazykových kurzech neziskových organizací.

Asi polovina respondentů pocházela z východoslovanského jazykového prostředí (celkem čtyři), zatímco zbytek z nich byli mluvčí angličtiny (celkem pět). Ve skupině byli studenti z různého sociálního prostředí, kteří zastávali různé profese, proto lze předpokládat, že i jejich zkušenost s psaním delších textů nejen v češtině, ale i v jiných jazycích, mohla být odlišná. Mezi respondenty ovšem podstatnou část tvořili lektori angličtiny (celkem čtyři).

### **6.2 Nasbíraná data**

Před samotným vstupem do kurzu měli studenti za úkol zpracovat jako pretest dvě písemné úlohy z modelových testů ke zkoušce CCE-B1. Jednalo se o (i) dopis kamarádovi a (ii) vyplnění dotazníku spokojenosti.

Během kurzu pak studenti napsali celkem čtyři další texty, přičemž se vždy jednalo o dvě verze téže úlohy. To znamená, že během prvního víkendu odevzdali dva dopisy a dva dotazníky. První verze dopisů byly podrobeny pedagogickému hodnocení, zatímco první verze dotazníků si studenti hodnotili vzájemně. Druhé verze dopisů i dotazníků obsahují případné implementace úprav navržených při hodnocení.

Vedle zmíněných písemných prací psali studenti v rámci kurzu během části experimentu zaměřené na získání dat o vzájemném hodnocení také komentáře ke třem textům svých kolegů. Po skončení kurzu pak měli za úkol vyplnit dotazník, v němž reflektovali své dojmy z výuky a hodnotili jednotlivé aktivity. Dotazník vyplnilo osm z devíti respondentů.

V experimentu jsem tedy k další analýze získala celkem 54 písemných prací, 27 „recenzí“ (vzájemných hodnocení) a odpovědi z osmi dotazníků (přílohy č. 3–7). Kompletní přehled nasbíraných dat ukazuje tabulka 3.

**Tabulka 3: Přehled nasbíraných dat**

aktivita	text typu dopis	text typu dotazník	hodnotící komentáře	závěrečný dotazník
pretest	9	9		
<b>1. víkend</b>				
psaní textu k pedagogickému hodnocení	9			
psaní druhé verze textu	9			
<b>2. víkend</b>				
psaní textu ke vzájemnému hodnocení		9		
vzájemné hodnocení („recenze“)			27	
psaní druhé verze textu		9		
reflexe kurzu				8

### 6.2.1 Zpracování dat z textů

Všechny získané texty včetně komentářů jsem převedla do textové podoby a přiřadila k nim čísla respondentů. Texty získané během víkendových lekcí jsem navíc podrobila důkladnější analýze, kterou přesněji popisuji v pododdílech níže.

#### Analýza a anotace dat ze vzájemného hodnocení

Komentáře ze vzájemného hodnocení jsem podrobila **obsahové analýze** z hlediska toho, jaké kategorie jevů se týkaly, jestli případně správně identifikovaly chyby nebo pozitivní jevy v textu, jestli nabízely návrhy úprav a nakolik byly tyto návrhy správné.

Jednotlivé jevy jsem kategorizovala na následujících rovinách: (i) téma a typ textu, dodržení zadání, (ii) strukturování textu, (iii) argumentace, (iv) rukopis, (v) pravopis a hláskosloví, (vi) gramatika, (vii) lexikum, (viii) slovosled, (ix) větná struktura, (x) interpunkce a (xi) formulace. Pokud se v rámci určité jednotky objevily chyby na více textových rovinách, identifikovala jsem je jako více chyb (např. když mělo jedno slovo nesprávnou gramatickou koncovku a navíc se v něm objevila také pravopisná chyba).

Identifikaci i návrh řešení jsem kategorizovala jako správné, částečně správné a nesprávné. Částečná správnost u identifikace znamená, že hodnotitelovo povšimnutí bylo jen částečně



přesné (identifikoval např. na rovině pravopisu a hláskosloví nesprávné *e* ve slově *vysvětlilá* (správně *vysvětlila*), ale už ne kvantitu *í*, *á*), nebo že jsem relevanci komentáře považovala za spornou. Částečná správnost u návrhu řešení pak znamená jen částečné vylepšení daného problému.

U každého návrhu úpravy jsem dále sledovala, nakolik ho hodnocený autor následně implementoval do svého textu. V tomto případě jsem rozlišovala implementace plně realizované, částečně realizované a nerealizované. Částečně realizované implementace odpovídají částečnému přijetí návrhu, u nerealizovaných návrhů jsem v poznámce sledovala, zda byl identifikovaný problém v druhé verzi textu vyřešen jiným způsobem (vynecháním části textu, přeformulováním). Pokud nebylo z povahy dat možné vyhodnotit kteroukoliv z výše uvedených kategorií, označila jsem ji v tabulce jako „x“ (nelze určit). Kompletní obsahovou analýzu komentářů ze vzájemného hodnocení uvádím v příloze č. 9.

Současně jsem u textů určených pro vzájemné hodnocení (tzn. u prvních verzí dotazníků) přistoupila k **chybové analýze** textů, v níž jsem chyby rozdělila do stejných kategorií jako u obsahové analýzy komentářů. U každé z chyb jsem pak sledovala, jestli si jí jednotliví hodnotitelé povšimli. Vzhledem k tomu, že v některých případech docházelo k tomu, že si respondenti nevšimli kompletní chybné pasáže, ale pouze jejího fragmentu nebo dílčího problému, byly v tabulce některé jevy, které logicky představují pouze jednu chybu, rozděleny do více řádků a počet chyb se tak navýšil. Rozdělení chyb do více řádků se promítlo i do obsahové analýzy prezentované výše.

Aby bylo obě tabulky možné propojit a identifikovat případné duplicity (tzn. místa v textu, jimiž se zabývalo více hodnotitelů), přiřadila jsem každé chybě identifikační číslo.

### **Analýza a anotace dat z pedagogického hodnocení**

Zpětná vazba poskytnutá k prvním verzím dopisů v rámci pedagogického hodnocení byla rovněž podrobena **obsahové analýze**. Zatímco u zpětné vazby ze vzájemného hodnocení byl sledován i obsah komentářů, které text hodnotily také jako celek, u pedagogického hodnocení byly obsahově analyzovány pouze komentáře zaměřující se na dílčí jevy, protože u nich bylo následně snazší sledovat případnou implementaci návrhů úprav. Kompletní obsahovou analýzu komentářů ze vzájemného hodnocení uvádím v příloze č. 10.

### **Ohodnocení textů na základě SERR (2006)**

Celkový posun mezi úrovní jednotlivých textů, tzn. vstupní úroveň jednotlivých studentů, úroveň prvních verzí textů a úroveň přepracovaných úloh pak ohodnotila s přihlédnutím

k jednotlivým jazykovým rovinám ještě další lektorka.<sup>15</sup> To poskytla nezaujatý pohled na data, a přestože je její hodnocení subjektivní a pravděpodobně do určité míry nepřesné, poskytně v diskusi další informace o možných přínosech zkoumané metody. Kompletní hodnocení lektorky uvádím v příloze č. 12.

### **Kvantifikace analýz**

U anotovaných dat jsem sledovala s ohledem na formulované hypotézy různé kombinace charakteristik a jejich hodnot zjištěných v chybových a obsahových analýzách. Data jsem vyhodnocovala jak souhrnně, tak podle jednotlivých respondentů a hodnotitelů. Na tomto základě jsem spočítala základní kvantitativní charakteristiky (frekvence, součty, průměry a podíly), které dovolují formulovat souhrnná i dílčí zjištění o zkoumaném souboru.

Přestože formativní hodnocení z principu není zaměřené pouze na hledání chyb. V práci je asi největší prostor věnován právě výzkumu povšimnutí a identifikace nesprávných jazykových prostředků, správnosti navržených úprav a otázce jejich následné implementace ze strany studentů. To je dáno zejména tím, že tyto jevy lze nejpřesněji a nejspolehlivěji sledovat, kvantifikovat a vyhodnocovat.

### **6.2.2 Zpracování dotazníků**

Analýza uzavřených odpovědí ze závěrečného dotazníku je v práci prezentována ve formě sloupcových grafů. Otevřené otázky byly podrobeny obsahové analýze.

## **6.3 Výsledky ve vztahu k výzkumným otázkám**

### **6.3.1 Kvalita komentářů u vzájemného hodnocení**

#### **Obsah komentářů**

Z obsahové analýzy komentářů ze vzájemného hodnocení (kompletní data viz příloha č. 9) vyplývá, že většina zpětné vazby se týkala hlavně negativních jevů. Z celkového počtu 316 komentářů bylo negativních 270 (83 %) a pozitivních 46 (17 %).

Pozitivní komentáře se u většiny hodnotitelů vztahovaly zejména k hodnocení textu jako celku (38, tj. 82,6 % z celkového počtu pozitivních komentářů). Dílčí části textu ve větší míře u svých spolužáků komentoval pozitivně pouze respondent č. 7, který byl autorem šesti z osmi pozitivních komentářů, které nehodnotily text jako celek. V tomto ohledu tedy většina studentů kopírovala formát zpětné vazby nastavený během pedagogického hodnocení (viz příloha č. 4).

---

<sup>15</sup> O ohodnocení textů jsem požádala kolegyni T. Štochlovou.

Obsahově se pak celkově pozitivní hodnocení týkalo zejména roviny tématu textu, typu textu a dodržení zadání (17, tj. 37 %) a lexika (14, tj. 30,4 %).

Jak napovídá tabulka 4 níže, nejvíce komentářů z jejich celkového počtu se týkalo roviny gramatiky (celkem 119 komentářů, tj. 37,7 % z celkového počtu komentářů). Na druhém místě, avšak s velkým odstupem pak hodnotitelé řešili pravopis a hláskosloví (celkem 60 komentářů, tj. 19 %). Tyto údaje vyplývají z povahy zkoumaného materiálu, v němž byly nedostatky v oblasti gramatiky a pravopisu nejvíce frekventované (gramatické chyby byly zastoupeny v 37,2 %, pravopisné chyby v 38,1 %, viz příloha s chybovou analýzou). Osvojování gramatických struktur je při přípravě na zkoušku na úrovni B1 navíc často v centru zájmu, protože studentům tato oblast obvykle dělá problémy. Nedostatky např. v argumentaci jsou naproti tomu jednak hůře rozpoznatelné, jednak se pravděpodobně budou více objevovat v komplexnějších textech.<sup>16</sup>

**Tabulka 4: Komentáře podle jednotlivých rovin a podle typu hodnocení**

<b>rovina</b>	<b>všechna hodnocení</b>	<b>pozitivní hodnocení</b>	<b>negativní hodnocení</b>
argumentace	13	3	10
formulace	20	1	19
gramatika	119	3	116
interpunkce	3	0	3
jednoduchost vět	8	2	6
lexikum	38	14	24
pravopis a hláskosloví	60	0	60
rukopis	10	2	8
slovosled	13	0	13
strukturování textu	3	3	0
téma a typ textu, zadání	27	17	10
nelze určit	2	1	1
<b>celkem</b>	<b>316 (100 %)</b>	<b>46 (14,6 %)</b>	<b>270 (85,4 %)</b>

### **Míra povšimnutí chyb**

Tento oddíl se ve vztahu k hodnocení zabývá tzv. výtěžností („recall“),<sup>17</sup> tedy tím, kolika chyb z jejich celkového počtu<sup>18</sup> si studenti při vzájemném hodnocení všimli.

<sup>16</sup> V tomto ohledu by mohl být zajímavý např. výzkum vzájemného hodnocení písemných prací u vyšších úrovní a výstavbově komplexnějších textů.

<sup>17</sup> Pojem výtěžnost („recall“) převzatý z informatiky je obvykle využíván jako jeden z údajů při hodnocení kvality nástrojů textové analýzy. Udává poměr relevantních výsledků analýzy ke všem relevantním výsledkům v daném vzorku bez ohledu na to, jestli byly nástrojem identifikovány (Cvrček, 2017).

<sup>18</sup> Pojem „celkový počet chyb“ se vztahuje k analýze chyb v textech, kterou jsem provedla (příloha č. 11).

Výsledky této analýzy naznačují (tabulka 5), že míra povšimnutí chyb v hodnoceném textu z pohledu hodnotitelů (tzn. jaký podíl chyb odhalil hodnotitel z celkového počtu chyb ve třech hodnocených textech) dosahovala při vzájemném hodnocení průměrně asi 26,1 %. Vzhledem k tomu, že na každém textu pracovali tři hodnotitelé, dostali ale jednotliví autoři v průměru alespoň jeden komentář k 52,8 % chyb (tabulka 6). To může sice vést k předpokladu, že větší počet hodnotitelů jednoho textu kompenzoval nízkou výtěžnost, jak ale ukázala analýza duplicit (podrobněji v oddílu 0), v případech, kdy proti sobě stojí protichůdné návrhy úprav, je pro studenty náročné zorientovat se v tom, který z nich je správný.

Při pohledu na tabulku týkající se jednotlivých hodnotitelů (tabulka 5) se také ukazuje, že mezi jednotlivými studenty byly při identifikaci chyb velké rozdíly a že někteří z nich prakticky vůbec nepřistoupili ke komentování jednotlivých jevů.<sup>19</sup> V tomto ohledu je zajímavé, že tento problém se může týkat také studentů, které lektor vyhodnotil jako relativně pokročilejší. Např. respondentku/hodnotitelku<sup>20</sup> č. 4 jsem sice zařadila mezi studenty se spíše vyšší úrovní jazykového projevu a předpokládala jsem, že bude kolegům poskytovat zpětnou vazbu vyšší kvality, vzhledem k nízké míře povšimnutí (nebo vědomého komentování) chyb ale respondenti č. 2 a 5, které hodnotila,<sup>21</sup> dostali výrazně méně podnětů ve vztahu k chybám v textu (tabulka 6). To může ukazovat na to, že student s vyšší jazykovou kompetencí v písemném projevu nemusí být stejně úspěšný při vyhledávání chyb v (cizích) textech.<sup>22</sup>

Výsledky chybové analýzy potvrzují (vedle srovnání s pedagogickým hodnocením, které uvedu dále) hypotézu, že respondenti dostanou více podnětů od lektora než od svých spolužáků. Zároveň byla potvrzena i hypotéza, že studenti neurčí správně všechny nedostatky v textech svých kolegů.

---

<sup>19</sup> Zde pochopitelně nelze zcela vyloučit jako možnou příčinu obtíže při práci s používanou technologií. Tento problém jsem se ale v kurzu snažila minimalizovat společným nácvikem vkládání komentářů u cvičného textu a průběžným monitorováním studentů během jejich práce.

<sup>20</sup> Číslování hodnotitelů v tabulce 4 odpovídá číslování respondentů v tabulce 5.

<sup>21</sup> Hodnotitel č. 4 hodnotil také respondenta č. 8, který však získal dostatečné množství komentářů díky shodě okolností: hodnotitelé č. 7 a 9, kteří byli jinak podprůměrní, každý identifikovali nejvíce chyb právě u tohoto respondenta.

<sup>22</sup> Zde je na místě podotknout, že vyhledávání chyb v textech je specifická dovednost. Tato aktivita může pro studenty představovat velice umělou a nezvyklou situaci, jednotliví studenti mohou mít také rozdílnou potřebu nácviku. Vliv může mít ale i nesprávný odhad lektora. U respondentky č. 4 jsem totiž předpokládala kvalitní zpětnou vazbu do velké míry na základě ústní interakce během diskusí v kurzu, kdy se projevovala jako velmi aktivní a úspěšně komentovala modelové texty. V jejím písemném projevu byly ovšem chyby v pravopisu a gramatice poměrně frekventované, a zřejmě jsem tedy mohla očekávat, že právě tyto nejčastější chyby pro ni bude nejnáročnější rozeznávat i v textech kolegů.

**Tabulka 5: Míra povšimnutí chyb podle hodnotitelů**

hodnotitel	celkem chyb k identifikaci	identifikováno chyb	podíl identifikovaných
Hodnotitel č. 1	80	33	41,3 %
Hodnotitel č. 2	73	27	37,0 %
Hodnotitel č. 3	90	44	48,9 %
Hodnotitel č. 4	56	1	1,8 %
Hodnotitel č. 5	73	25	34,2 %
Hodnotitel č. 6	77	11	14,3 %
Hodnotitel č. 7	56	15	26,8 %
Hodnotitel č. 8	112	19	17,0 %
Hodnotitel č. 9	52	7	13,5 %
<i>průměr hodnotitelů</i>			26,1 %
<b>celkem</b>	<b>669</b>	<b>182</b>	<b>27,2 %</b>

**Tabulka 6: Míra povšimnutí chyb podle respondentů**

respondent	celkem chyb v textu respondenta	chyby identifikované aspoň jednou	podíl identifikovaných
Respondent č. 1	26	15	57,7 %
Respondent č. 2	12	4	33,3 %
Respondent č. 3	26	12	46,2 %
Respondent č. 4	25	14	56,0 %
Respondent č. 5	18	3	16,7 %
Respondent č. 6	42	31	73,8 %
Respondent č. 7	22	15	68,2 %
Respondent č. 8	26	13	50,0 %
Respondent č. 9	26	19	73,1 %
<i>průměr respondentů</i>			52,8 %
<b>celkem</b>	<b>223</b>	<b>126</b>	<b>56,5 %</b>

Na míru povšimnutí chyb u hodnotitelů i hodnocených jsem se doplňkově podívala také ve vztahu ke komentovaným rovinám. Tato data naznačují, že nejvyšší podíly chyb identifikovaných v textech respondentů alespoň jednou se objevily na rovinách formulací (11 ze 16, tj. 68,8 %) a gramatiky (51 z 83, tj. 61,4 %).

### Správnost identifikace a návrhů úprav

Z druhé strany jsem se podívala na správnost identifikace a správnost návrhů úprav ve smyslu tzv. „precision“<sup>23</sup> (tedy na to, jak velký podíl komentářů z jejich celkového počtu byl relevantní

<sup>23</sup> Pojem přesnost („precision“) převzatý z informatiky je obvykle využíván jako jeden z údajů při hodnocení kvality nástrojů textové analýzy. Je definován jako procentuální poměr relevantních výsledků analýzy ke všem výsledkům, které byly analýzou získány (Cvrček, 2017).

z těchto hledisek). V tomto případě se dívám na všechny komentáře, tedy jak pozitivní, tak negativní (tedy kolik povšimnutých chyb bylo hodnotiteli správně identifikováno).

Z hlediska správnosti identifikace studenti ve většině případů (v 85,1 %) publikovali komentář správný nebo alespoň částečně správný (viz levá část tabulky 7), přestože se jejich úspěšnost v tomto ohledu liší na jednotlivých rovinách. Ačkoli jsou naměřené absolutní hodnoty nízké, data naznačují, že hodnotitelé měli obtíže s identifikací chyb a vhodných prostředků na úrovni argumentace (zde možná hraje roli subjektivní vnímání logiky textu) a slovosledu.

Z hlediska návrhů úprav je podíl správných návrhů nižší (49,1 %), což je logické jednak s přihlédnutím ke složitosti úkolu, jednak z toho důvodu, že některé (zejména pozitivní) komentáře žádné návrhy úprav nevyvolávají. Podíl nesprávných návrhů je podobný podílu nesprávných identifikací.

Jak ukazuje pravá strana tabulky 7 nesprávné návrhy úprav respondenti poskytovali v absolutních číslech nejvíce na úrovni gramatiky. Celkový počet 22 nesprávných návrhů řešení odpovídá podílu 18,5 % ze všech návrhů řešení na této rovině, což je vysoký podíl ve srovnání s ostatními rovinami. Vyšší podíl nesprávných návrhů se objevil jen na rovině slovosledu (zde jsou však celkové absolutní počty nízké a nelze z toho mnoho vyvozovat).

Tato část dat tedy potvrdila hypotézu, že studenti nebudou k chybám vždy nabízet správný návrh úpravy.

**Tabulka 7: Identifikace a návrh řešení všech komentářů podle správnosti**

rovina	identifikace				navržené řešení			
	správná	část. spr.	nespr.	nelze u.	správné	část. spr.	nespr.	nelze u.
argumentace	7	0	6	0	0	0	2	11
formulace	16	1	2	1	14	3	2	1
gramatika	98	8	11	2	66	14	22	17
interpunkce	3	0	0	0	3	0	0	0
jednoduchost vět	4	3	1	0	3	0	0	5
lexikum	28	2	5	3	9	3	4	22
pravopis a hláskosloví	51	4	5	0	43	8	5	4
rukopis	10	0	0	0	6	0	0	4
slovosled	9	1	3	0	9	0	3	1
strukturování textu	3	0	0	0	0	0	0	3
téma a typ textu, zadání	20	1	2	4	2	1	0	24
nelze určit	0	0	1	1	0	0	0	2
<b>celkem 316</b>	<b>249</b>	<b>20</b>	<b>36</b>	<b>11</b>	<b>155</b>	<b>29</b>	<b>38</b>	<b>94</b>
<b>v % (100)</b>	<b>(78,8)</b>	<b>(6,3)</b>	<b>(11,4)</b>	<b>(3,4)</b>	<b>(49,1)</b>	<b>(9,2)</b>	<b>(12)</b>	<b>(29,7)</b>

Následně jsem se podívala jen na komentáře s negativním hodnocením, u nichž respondenti správně identifikovali nedostatky a zároveň navrhli nějaké řešení. K většině správně identifikovaných nedostatků respondenti nabízeli správná nebo částečně správná řešení, ačkoli povětšinou formou pouhé opravy nebo návrhu. Tabulka 8 shrnuje počty správných, částečně správných a nesprávných návrhů řešení ve vztahu k obsahu komentáře. Analogickou tabulku jsem sestavila pro pedagogické hodnocení textů z prvního víkendu. Z porovnání vyplývá, že studenti na rozdíl od lektora přistupovali výrazně méně často k poskytování vysvětlení (tabulka 9). Tento rozdíl je samozřejmě očekávatelný vzhledem k tomu, že jde o aktivitu, která je náročná jak časově, tak kognitivně. Nelze navíc očekávat, že studenti cizího jazyka budou vykazovat např. přesnou orientaci v pravidlech, schopnost hlouběji analyzovat jednotlivé jevy nebo i explicitní znalost gramatických pouček, a to i v případech, kdy správně identifikují místo, na němž by bylo vhodné použít jinou formu nebo jazykový prostředek.

**Tabulka 8: Vztah správnosti navrženého řešení a obsahu komentáře u správně identifikovaných nedostatků**

navržené řešení	celkem	oprava, návrh	vysvětlení + návrh, vysvětlení	jiné
správné	148	129	14	5
částečně správné	22	18	4	0
nesprávné	12	12	0	0
<b>celkem</b>	<b>182</b>	<b>159</b>	<b>18</b>	<b>5</b>

**Tabulka 9: Obsah komentářů pedagogického hodnocení**

	celkem	oprava, návrh	vysvětlení + návrh, vysvětlení	jiné
navržené řešení	294	149	142	3

Počet komentářů s vysvětlením je u vzájemného hodnocení natolik nízký, že není možné porovnávat poměr, v němž studenti realizovali návrhy úprav, které obsahovaly vysvětlení, oproti těm, jež ho neobsahovaly.

### 6.3.2 Implementace úprav

#### Souhrnná analýza implementace komentářů

Míru implementace návrhů úprav ze vzájemného hodnocení s ohledem na jejich správnost shrnuje tabulka 10. Tabulka zahrnuje pouze případy, kdy bylo možné určit správnost navržených řešení a zároveň vyhodnotit naplněnost realizace návrhu.<sup>24</sup>

Celkové počty přijatých návrhů ukazují, že respondenti přijali méně než polovinu návrhů od hodnotitelů (z 221 návrhů plně realizovali 99 implementací, zatímco jich nerealizovali 107). Rozbor duplicitních návrhů (viz dále) ukazuje, že respondenti ve skutečnosti plně realizovali u 176 jevů 75 implementací a nerealizovali jich 79). Počet nerealizovaných návrhů tak nezávisle na způsobu výpočtu ukazuje na určitou nedůvěru k úsudku kolegů. Tu ostatně naznačuje i srovnání s návrhy úprav z pedagogického hodnocení. Tabulka 11 ukazuje, že studenti přijali celkem 154 návrhů úprav od lektora z celkových 289. Podíl se sice může zdát velmi podobný tomu u vzájemného hodnocení, bližší analýza však ukázala, že ze 125 nerealizovaných úprav se jich 102 týkalo pasáží, které studenti v druhých verzích textu zcela vynechali nebo přeformulovali.<sup>25</sup> Ve vzájemném hodnocení bylo případů vynechání nebo přeformulování komentované pasáže pouze 28 ze 79 nerealizovaných implementací (po zohlednění duplicitních komentářů). To znamená, že 51 (29 %) míst v textu, k nimž dostali respondenti alespoň jeden návrh úpravy, nechali jejich autoři zcela beze změny. U pedagogického hodnocení se tento jev týkal 8 % míst. Výsledky tak podporují hypotézu, že studenti provedou méně úprav na základě komentářů od kolegů než na základě pedagogického hodnocení.

---

<sup>24</sup> V tabulce 9 jsou zahrnuty z pohledu hodnotitelů jak komentáře vztahující se k textu jako k celku, tak případy, kdy hodnocení dostali k jednomu jevu více návrhů řešení. Tento problém detailněji rozeberu v následujících odstavcích, kde jsou implementace komentářů rozebrány z hlediska respondentů.

<sup>25</sup> Tuto strategii ve vysoké míře zvolili respondenti č. 6, 7 a 9, kteří přistoupili k výrazně větším změnám v textu, než které jim byly doporučeny v rámci zpětné vazby (srov. příloha č. 4, str. 7–8, 10 a příloha č. 6, str. 3, 4).



**Tabulka 10: Míra implementace návrhů na jednotlivých rovinách podle správnosti (vzájemné hodnocení)**

navržené řešení realizace	správné			nesprávné			částečně správné		
	plně real.	část. real.	nereal.	plně real.	část. real.	nereal.	plně real.	část. real.	nereal.
argumentace	0	0	0	2	0	0	0	0	0
formulace	2	3	9	0	0	2	0	0	3
gramatika	37	3	26	6	1	14	4	2	8
interpunkce	3	0	0	0	0	0	0	0	0
jednoduchost vět	0	0	3	0	0	0	0	0	0
lexikum	3	1	5	1	0	3	0	0	3
pravopis a hláskosloví	26	0	17	2	0	3	6	2	0
rukopis	1	0	5	0	0	0	0	0	0
slovosled	5	2	2	1	0	2	0	0	0
strukturování textu	0	0	0	0	0	0	0	0	0
téma a typ textu, zadání	0	0	2	0	0	0	0	1	0
nelze určit	0	0	0	0	0	0	0	0	0
<b>celkem (221)</b>	<b>77</b>	<b>9</b>	<b>69</b>	<b>12</b>	<b>1</b>	<b>24</b>	<b>10</b>	<b>5</b>	<b>14</b>
<b>celkem bez ohledu na správnost</b>	<b>plně real.</b>	<b>část. real.</b>	<b>nereal.</b>						
	99	15	107						

**Tabulka 11: Implementace oprav v pedagogickém hodnocení podle rovin**

	plně realizovaná	částečně realizovaná	nerealizovaná
argumentace	2	1	1
formulace	6	2	7
gramatika	64	2	34
interpunkce	6	1	6
jednoduchost vět	4	0	0
lexikum	17	2	17
pravopis a hláskosloví	39	1	48
rukopis	0	0	1
slovosled	16	1	10
strukturování textu	0	0	1
téma a typ textu, zadání	0	0	0
nelze určit	0	0	0
<b>celkem (289)</b>	<b>154</b>	<b>10</b>	<b>125</b>

Respondenti u vzájemného hodnocení celkově častěji přijímali úpravy na rovině pravopisu a hláskosloví než na rovině gramatiky (viz tabulka 10 výše). Jako zajímavá z hlediska přijímání úprav od spolužáků se jeví neochota přijímat úpravy z kategorie formulací. Z těch studenti realizovali minimum, přičemž možným důvodem této neochoty může být např. určitá nejednoznačnost této kategorie chyb a návrhů úprav. Často totiž může chybět uspokojivé vysvětlení, protože hodnotitelé se odkazují v podstatě pouze na svůj jazykový cit, viz (1),

v literatuře pak některé studie reflektují určitý vliv autorství (Chang, 2016, s. 103). K přesnému zjištění motivů za těmito rozhodnutími by ale bylo zapotřebí sběru dalších dat, např. následné rozhovory se studenty.

(1) Odvolání se na jazykový cit

H 1: *To zní trošku divny. Lepší by mohl „Pomohla s mé češtinou, kterou je velmi slabou“*

#### **Analýza povšimnutých jevů s přihlédnutím k duplicitním komentářům (vzájemné hodnocení)**

V předchozí části textu jsem uvedla, že některé jevy byly komentovány více hodnotiteli, v tomto oddílu se proto blíže zaměřuji na případy, kdy se komentáře více hodnotitelů k jednomu jevu překrývaly.

Na celkových 176 jevů, které studenti během vzájemného hodnocení v textech komentovali, připadá celkem 129 unikátních návrhů (tzn. že k příslušnému jevu další hodnotitelé návrh řešení nenabídli) a 47 duplicitních návrhů (tzn. jevů, k nimž poskytlo nějaký návrh řešení více hodnotitelů).<sup>26</sup> Jejich rozprostření na jednotlivých rovinách textu blíže popisuje tabulka 10.

**Tabulka 12: Jevy s unikátními a duplicitními návrhy oprav podle rovin**

	celkem jevů	jevů s unikátními návrhy	jevů s duplicitními i návrhy
argumentace	6	5	1
formulace	16	13	3
gramatika	68	42	26
interpunkce	2	1	1
jednoduchost vět	2	2	0
lexikum	16	12	4
pravopis a hláskosloví	44	32	12
rukopis	6	6	0
slovosled	11	11	0
strukturování textu	0	0	0
téma a typ textu, zadání	4	4	0
nelze určit	1	1	0
<b>celkem</b>	<b>176</b>	<b>129</b>	<b>47</b>

Vytvoření tabulky 12 umožnila chybová analýza textů spolu s přiřazením identifikačního čísla jednotlivým chybám. Díky tomuto postupu je možné také detailněji sledovat, co se stane v případě, kdy respondenti dostanou k jednomu jevu více podnětů. Jak naznačuje příklad (2)

<sup>26</sup> V 39 případech obdrželi respondenti k jednomu jevu dva návrhy, v osmi případech byly návrhy tři. V trojnásobných návrzích se neobjevila více než dvě různá řešení a jejich míra implementace je zároveň podobná návrhům duplicitním. Pro jednoduchost jsem je proto zahrnula mezi návrhy duplicitní (a tímto termínem je také označuji).

níže, někdy je pro studenta komplikované si ze tří různých návrhů úprav zvolit ten nejsprávnější, někdy v případě nejistoty volí taktiku vyhýbání a přeformuluje text tak, aby z něj danou pasáž mohl zcela vynechat, viz (3).

(2) Přijetí méně kvalitního návrhu

R 6.1: *Pomáhala mě se moje čeština, který je velmi slabý.*

H 1: *To zní trošku divný. Lepší by mohl „Pomohla s mé češtinou, kterou je velmi slabou“*

H 8: *lepší, pomáhala mě češtině, který byla velmi slábý nebo špatné.*

H 3: *Pomáhala mě s moje češtinou, která je velmi slabá.*

R 6.2: *Pomohla se mé češtinou, kterou je velmi slabou.*

(3) Taktika vyhýbání

R 6.1: *[Měla jsem radost s toho, že paní učitelka byla s námu laskavá a v všem nám vysvětlila, ráda poslouchala] nás [a odpovídla na všechny naše otázky.]*

H 1: ?

R 6.2: *Měla jsem radost z toho že jsme pracovali ve skupině a jsme měli možnost seznámit se s ostatními studentům.*

Tabulka 13 dále shrnuje počty duplicitních úprav na jednotlivých úrovních ve vztahu k počtu komentovaných jevů. Nejčastěji se duplicitní komentáře k textům spolužáků objevovaly v rovině gramatiky. Z celkového počtu 102 komentářů, které se vztahovaly k celkem 47 jevům v textu jich do této kategorie spadalo celkem 59.

**Tabulka 13: Duplicitní návrhy oprav ve vztahu k jevům podle rovin**

	celkem komentářů	jevů v textu
argumentace	3	1
formulace	6	3
gramatika	59	26
interpunkce	2	1
lexikum	8	4
pravopis a hláskosloví	24	12
<b>celkem</b>	<b>102</b>	<b>47</b>

Tabulky 14 a 15 pak detailněji shrnují chování hodnocených autorů, pokud jde o realizace návrhů řešení. Zatímco v tabulce č. 14 jsou zahrnuty duplicitní komentáře, které se spolu shodují, v tabulce č. 15 jsou komentáře, které se od sebe liší v návrzích úprav.

Zatímco u 28 jevů z celkových 47 komentovaných se návrhy hodnotitelů shodovaly, u 19 jevů nabízeli jednotliví respondenti spolužákům různá řešení. I přes nízké absolutní počty se ukazuje tendence, že v případě návrhů, které se shodovaly, studenti přistupovali k implementaci úprav častěji než v případě neshodujících se návrhů. Ve chvíli, kdy se od sebe návrhy lišily, studenti návrhy implementovali nejvíce v případech, kdy proti sobě stál správný a částečně

správný návrh – tedy v případě, že se od sebe lišily nejméně. Toto chování opět naznačuje nejistotu při rozhodování o správnosti opravy, stejně jako nerozhodnost ohledně toho, jestli může autor hodnotiteli důvěřovat.

**Tabulka 14: Implementace oprav jevů s totožnými duplicitními návrhy**

duplicitní návrhy	celkem jevů	implementace			vynechání / přeformul.
		plně realizovaná	částečně realizovaná	nerealizovaná	
správné	21	14	0	7	5
částečně správné	4	1	1	2	0
nesprávné	1	1	0	0	0
nelze určit	2	0	0	2	0
<b>celkem</b>	<b>28</b>	<b>16</b>	<b>1</b>	<b>11</b>	<b>5</b>

**Tabulka 15: Implementace oprav jevů s odlišnými duplicitními návrhy**

duplicitní návrhy	implementace		konkur. návrhy	implementace		nerealiz. vynech. / přeform.
	plně realiz.	částečně realiz.		plně realiz.	částečně realiz.	
u 7 jevů správné vs. nesprávné:						
správné	2	0	nesprávné	2	1	2
u 7 jevů správné vs. částečně správné:						
správné	4	1	část. spr.	1 (+3) <sup>27</sup>	0	1
u 2 jevů částečně správné vs. nesprávné:						
část. spr.	0	0	nesprávné	1	0	1
u 3 jevů správné vs. správné alternativní:						
správné	2	1	správné alt.	-	-	0
<b>celkem 19 jevů</b>	<b>8</b>	<b>2</b>		<b>4</b>	<b>1</b>	<b>4</b>

Jako doplnění pro úplnou systematizaci dostupných informací pak může sloužit ještě tabulka 16, která shrnuje data k unikátním návrhům řešení. Z tabulky vyplývá, že studenti v případě, kdy dostali pouze jeden návrh řešení realizovali jen asi polovinu úprav. S přihlédnutím k tabulkám 14 a 15 lze tedy konstatovat, že autoři k úpravám přistupovali častěji v případech, kdy se podporovala stanoviska různých hodnotitelů. Díky tabulkám 14–16 je také nakonec možné souhrnně určit, že z celkového počtu 176 jevů s návrhy oprav bylo 76 implementací realizováno plně, 9 částečně, zatímco nebylo realizováno 79 implementací a v 12 dalších případech nebylo možné implementaci návrhu určit. Ve 28 případech nerealizovaných návrhů byl text vynechán nebo přeformulován.

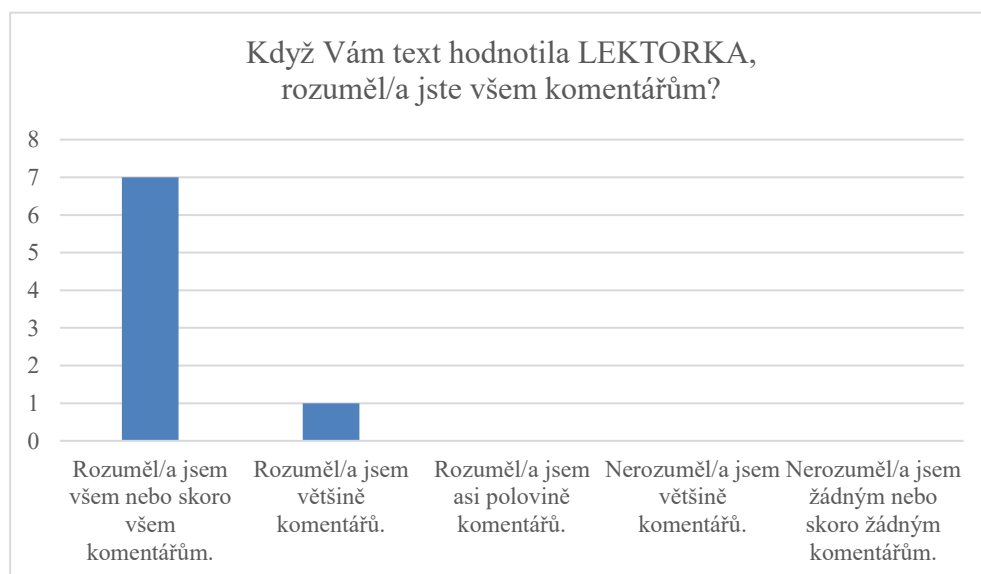
<sup>27</sup> Částečné úpravy, které jsou zahrnuty pod správnou úpravou.

**Tabulka 16: Implementace oprav jevů s unikátními návrhy**

unikátní návrhy	celkem jevů	implementace			vynechání / přeformul.
		plně realizovaná	částečně realizovaná	nerealizovaná	
správné	81	38	3	40	12
částečně správné	25	4	2	5	0
nesprávné	11	6	0	19	4
nelze určit	12	-	-	-	3
<b>celkem</b>	<b>129</b>	<b>48</b>	<b>5</b>	<b>64</b>	<b>19</b>

### 6.3.3 Vnímání vzájemného hodnocení studenty

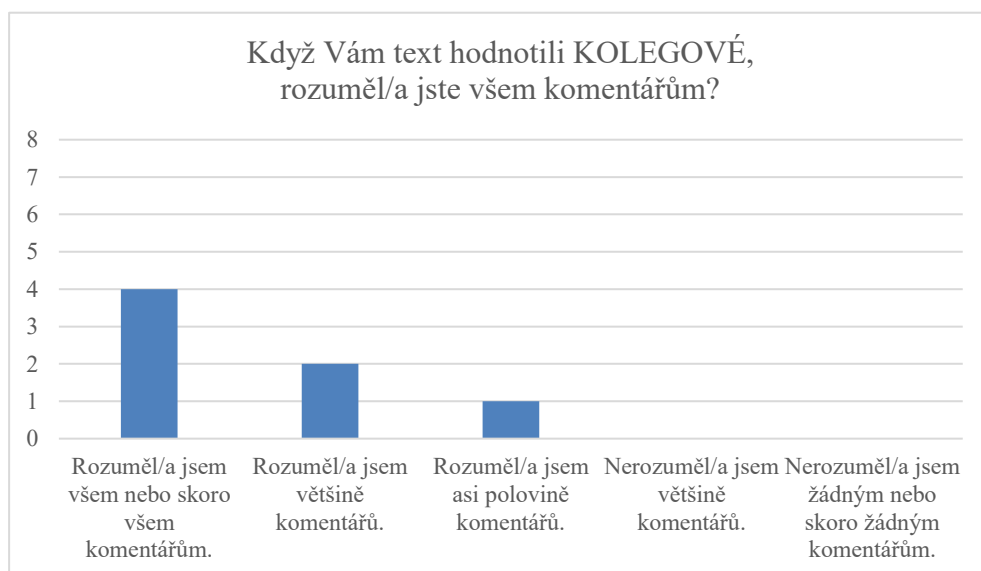
Z analýzy odpovědí v dotazníku, který respondenti vyplnili po skončení kurzu, vyplývá, že většina studentů hodnotí zpětnou vazbu od lektora jako srozumitelnější a užitečnější než zpětnou vazbu od spolužáků (grafy 1–4). Hypotéza o subjektivním hodnocení užitečnosti se tedy potvrdila.

**Graf 1**

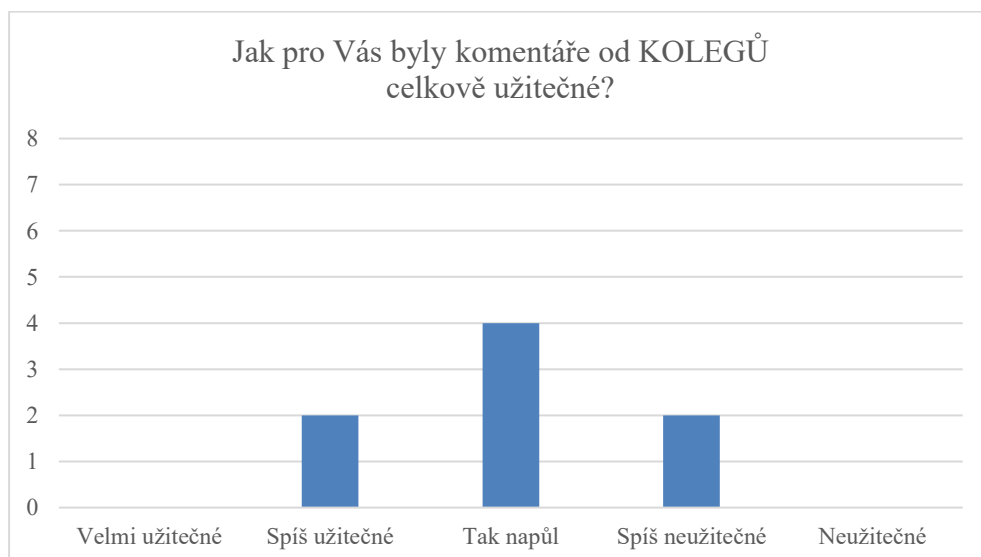
**Graf 2**



**Graf 3**



**Graf 4**



Odpovědi na otevřené otázky naznačují, že pro studenty byla klíčová spolehlivost zpětné vazby, pokud jde o správnou identifikaci problematických míst a návrhy úprav. Zatímco komentáře od učitele popisují studenti často jako „odborné“ (4 a–c), u komentářů od spolužáků hodnotili negativně jejich nevyrovnanou kvalitu a z ní plynoucí nejistotu týkající se jejich správnosti (5 a–c).

- (4) Jaké výhody a nevýhody podle Vás mělo, když Vám text komentovala LEKTORKA?
- a. *Jak jsem naspala, pro mě korektura z učitelky byla mnohem víc lepší než od kolegů. Učitel je rodilý mluvčí, je odborník(a), je schopný vidět celé našich chyb, je schopný napsat správné fráze anebo navrhnout další možnosti, obvykle rozumí proč studenty dělá chyby a to by mohl používat v následních cvičeních, aby studenty naučili něco správného od učitele.*
  - b. *Výhody: byli okomentované všechny naše chyby; typi jak to můžeme vylepšit a proč.*
  - c. *Víme, že je to 100% správné.*
- (5) Jaké výhody a nevýhody podle Vás mělo, když Vám text komentovali KOLEGOVÉ?
- a. *V komentáře svých kolegů jsem moc nevěřil. Raději bych měl kritiku od lektorky.*
  - b. *Nejsem jistá když moje kolegové dělají správný komentář*
  - c. *Jak jsem naspala, pro mě korektura z učitele je mnohem víc lepší než od kolegů. Studenty nejsou rodilý mluvčí, nejsou odborník(a), často nejsou schopné vidět chyby, často nejsou schopné napsat nebo neví správné fráze anebo navrhnout další možnosti/synonymy, a často nerozumí, co kolegy chci říct. Na druhé straně, když jsou na stejné úrovni, pomohli by vzájemně, protože mají podobné (ale ne vždycky stejné fráze) a jejich nápady by mohli být docela užitečné, ale vždycky potřebují pomoc učitele.*

Pokud jde o hodnocení přínosů vzájemného hodnocení oproti hodnocení pedagoga, dotazník nepotvrdil hypotézu, že studenti budou kladně hodnotit větší množství komentářů. Z otevřených odpovědí lze usuzovat, že důvodem byla nevyrovnaná kvalita zpětné vazby a z ní plynoucí nedůvěra ke komentářům spolužáků, z nichž si museli studenti vybírat (5 a–c; 6).

- (6) Co byste na kurzu změnil/a?
- Myslím si, že v druhé části kurzu, když jsme dělali komentáře k vyplněným dotazníkům sveh kolegů-cizinců, nebyli naši komentáře moc užiteční. Proto mohli bysme udělat ti komentáře, jako cvičení, ale pak paní lektorka měla by je zkontrolovat, jestli jsou správné ohledně češtiny, dát nam svůj feedback, a jenom pak je odesílat dotyčným kolegům. Jinak já nebyla jsem si jistá, jestli někteří komentáře odpovídají pravidlům českého jazyka a jestli vůbec je vnímat.*

Hypotéza týkající se pozitivního přijetí možnosti vidět texty spolužáků se spíše potvrdila (Graf 5). Studenti sice psaní komentářů popisovali v dotazníku jako velmi náročnou aktivitu, během níž se potýkali s pocitem, že nemají v češtině dostatečné kompetence, a vyjádřili potřebu

většího nácviku (7 a–c), zároveň ale někteří z nich oceňovali možnost inspirovat se nápady a jazykem spolužáků a učit se z jejich chyb (7 c, d).

**Graf 5**



- (7) Jaké pro Vás tato zkušenost měla výhody a nevýhody?
- Tato zkušenost pro mě nebyla moc dobře. Neměla jsem z toho zkušenost v češtině, nejsem odborník a to bylo moc naročný. Potřebovala bych víc vyučování, směrnic, fráz a podpory učitelky.*
  - Byla to úplně nová zkušenost, psát komentáře bylo těžké, jelikož nejsem si jistá ve svých znalostech, vím, že jsem jenom na začátku, a musím se ještě hodně učit.*
  - Byl zajímavý vidět jiné texty a nápady od kolegy. Byl 'těžký' dělat komentář se textem který jsou mnoho lepší než moje..*
  - Na chybách druhých jsem se naučila psát i dopis a dotazník.*

V odpovědích, které hodnotily kurz jako celek, se navíc objevily také postřehy, které kladně hodnotily možnost učit se z chyb někoho jiného a popisovaly pocit vyššího sebevědomí (8 a–c). Tyto kladné komentáře se ovšem nemusejí vztahovat přímo k fázi, v níž studenti hodnotili texty ostatních, ale mohou souviset např. se společným vyvozováním kritérií na základě úspěšných a neúspěšných textů. Naopak negativně bylo v těchto otázkách opakovaně hodnoceno množství studentů ve skupině a také jejich nevyrovnaná úroveň (9 a–c).

- (8) Co se Vám na kurzu líbilo?
- Libilo se mi, že jsme analizovali texty jiných kolegů-cizinců a viděli jsme nejčastější chyby. Také jsme probrali pravidla psaní neformálního dopisu a vyplnění dotazníku. Paní lektorka odborně odpovídala na otázky ohledně češtiny, což já si osobně moc vážím. Udělali jsme zásobu synonymů pro nejvíc potřebné slova v textech.*



- b. *Učitelka byla profesionální; naučila jsem se něco nového (syntax (byt/se), nové/důležité informace na konci vět, synonymy, měla jsem možnost vidět příklady dalších studentů*
  - c. *Dodalo mi to sebevědomí a ukázalo mi, na čem musím pracovat*
- (9) Jaké výhody a nevýhody podle Vás mělo, když Vám text komentovali KOLEGOVÉ?
- a. *Pro mě lepší když jenom par studentů bude na kurzy..*
  - b. *Asi to bylo by lepší, kdyby trochu méně studentů.*
  - c. *Aby byli všichni studenti na úrovni B1*

## 6.4 Diskuse

Data naznačují, že studenti připravující se ke zkoušce na úrovni B1 vykazují nízkou míru spolehlivosti při povšimnutí chyb, většinu identifikovaných nedostatků ovšem dokážou správně určit a doplnit je o dostatečně kvalitní návrh úpravy.

Analýza implementace chyb naproti tomu ukázala, že studenti jsou často neochotní návrhy úprav realizovat. Tuto nedůvěru k úsudku kolegů pak potvrdilo i vyhodnocení následného dotazníkového šetření mezi respondenty, které ukázalo, že studenti ke zpětné vazbě od kolegů pocítují nedůvěru a považují ji za nespolehlivou.

Tato zjištění mohla být do velké míry dána složením skupiny, které bylo sice poměrně heterogenní, bližší pohled na texty u nich ovšem naznačuje, že v projevu studentů převažovaly chyby na úrovni gramatiky, pravopisu a hláskosloví, zatímco na dalších rovinách jejich texty často přesahovaly úroveň B1. Chyby v gramatice a v pravopisu pak převažovaly v textech všech studentů. Je tedy možné předpokládat, že ačkoli byli hodnotitelé při identifikaci chyb v gramatice nadprůměrně úspěšní, bylo pro ně v řadě případů obtížné nabídnout k identifikovanému jevu správný návrh úpravy. Na straně hodnocených pak pravděpodobně bylo náročné se v návrzích zorientovat a dělat rozhodnutí o jejich implementaci.

Hodnocení celkové úrovně textů, které provedla jiná lektorka, naznačuje u některých respondentů určitý posun v jazykové úrovni. Tento posun lze vysledovat jak mezi pretestem a první verzí textu, tak mezi prvními a druhými verzemi dopisů i dotazníků. To může do jisté míry nasvědčovat jednak tomu, že někteří studenti zlepšili svůj projev jak na základě zpětné vazby od pedagoga (posun celkem u sedmi textů), tak na základě zpětné vazby od kolegů (posun celkem u tří textů). Vzhledem k tomu, že se u druhých verzí úloh jednalo do velké míry o zanášení oprav do již existujícího textu, nelze s jistotou tvrdit, že by se efekt projevil i u psaní dopisů a dotazníků na jiná témata a že by tedy byl dlouhodobý. Posun mezi pretesty a prvními verzemi dotazníků a dopisů (zaznamenán celkem u pěti textů) by mohl naznačovat, že znalost kritérií zlepšuje kvalitu vypracování úlohy, zde ovšem nemůžeme vyloučit např. faktor aktuálního rozpoložení, lepšího zvládnutí nového tématu atp.

Sporná efektivita implementace vzájemného hodnocení v provedeném experimentu mohla být do velké míry dána charakterem jednorázového intenzivního kurzu, v němž studenti neměli větší prostor na nácvik ani větší podporu lektora. Ze zjištění odborné literatury prezentované v teoretické části totiž jasně vyplývá (podrobněji viz oddíl 4.1.2), že vzájemné hodnocení má u studentů větší úspěch, pokud je doplněné o zpětnou vazbu od lektora a pokud je v hodinách velký prostor věnován jeho nácviku. Domnívám se tedy, že další výzkumy vzájemného hodnocení by bylo vhodné podniknout v rámci dlouhodobých kurzů, kde by tyto zásady bylo možné dodržet lépe.

Je také namístě podotknout, že ačkoli může výzkum na dané skupině studentů naznačovat určité tendence, pro jejich ověření by bylo zapotřebí zkoumat výrazně větší vzorek, jenž by umožnil za použití testů inferenční statistiky zobecnit data i nad rámec zkoumané skupiny.

## **6.5 Doporučení pro lektory**

Využití vzájemného hodnocení v přípravě studentů k písemné části zkoušky na úrovni B1 se v podniknutém experimentu ukázalo spíše jako neefektivní. Z hlediska výukového procesu by se při implementaci vzájemného hodnocení měli lektori důsledně zaměřit na provedení dílčích změn, které by podle mého názoru jeho kvalitu jednak vylepšily, jednak by vedly k pozitivnějšímu hodnocení této aktivity ze strany studentů, a tedy pravděpodobně i k jejich větší motivaci. Konkrétní doporučení nastiňuji v následujících pododdílech.

### **6.5.1 Předběžné vyhodnocení vhodnosti využití v dané skupině**

Před samotným zvolením metody vzájemného hodnocení by lektor podle mého názoru měl analyzovat texty studentů z hlediska kategorie pozitivních a negativních jevů, které v nich převažují. Využití vzájemného hodnocení podle mě není vhodné u skupin, v nichž převažují studenti, kteří mají opakovaně problémy pouze s gramatikou, zatímco zbytek textu obvykle splňuje stanovená kritéria. Jak ukázal experiment, studenti v této oblasti vykazují velkou nejistotu: vedle nízké výtěžnosti chyb dochází také k nízké nebo zmatečné realizaci navržených úprav. Kromě toho se domnívám, že tato aktivita může v důsledku vést také k negativním jevům, které představuje např. fixace nesprávných tvarů.

Vzájemné hodnocení naproti tomu může být vhodnější např. ve skupinách, v nichž mají někteří studenti problém používat dostatečně širokou slovní zásobu, v nichž jim dělá problémy napsat text v požadovaném rozsahu anebo ho vhodně strukturovat. Domnívám se, že v takových skupinách studenti mohou těžit také z možnosti vidět texty jiných a inspirovat se například tím, jak jejich spolužáci efektivně pracují s osvojenými jazykovými prostředky.

### **6.5.2 Postupná implementace a nácvik**

Vzájemné hodnocení může pro studenty představovat časově i kognitivně náročnou aktivitu. Experiment, který jsem podnikla v rámci intenzivního kurzu, podle mého názoru jasně ukázal, že implementace této metody není efektivní ve chvíli, kdy je k ní přistoupeno jednorázově a bez důkladného nácviku. Aby byla metoda úspěšnější, doporučila bych poskytnout studentům více vzorových textů, na nichž si mohou formulaci zpětné vazby zkoušet, a tuto aktivitu v počátcích systematicky doplňovat o korekce ze strany lektora. Domnívám se, že postupné zavádění vzájemného hodnocení doplněné o velkou podporu lektora v počátečních fázích nácviku zlepší také vnímání této aktivity ze strany studentů, protože jednak budou mít možnost si na využití této metody zvyknout, jednak si také budou moci vybudovat dostatečnou sebedůvěru ve vlastní hodnotitelské dovednosti a získat větší důvěru ke zpětné vazbě od spolužáků.

### **6.5.3 Doplnění o zpětnou vazbu od lektora**

Jak naznačují výsledky studií prezentovaných v teoretické části i zjištění provedená na základě podniknutého experimentu, studenti mají tendenci negativně hodnotit využití vzájemného hodnocení, pokud není doplněno o zpětnou vazbu od lektora, kterou považují za odbornější a spolehlivější. Domnívám se tedy, že vzájemné hodnocení by nikdy nemělo ve výuce na dané jazykové úrovni plně nahrazovat hodnocení lektora, ale mělo by být využíváno spíše jako doplňkové cvičení.

## 7 Shrnutí a závěr

Práce si kladla za cíl na základě experimentu srovnávajícího pedagogické a vzájemné hodnocení ověřit objektivní využitelnost vzájemného hodnocení v přípravě studentů k písemné části zkoušek na úrovni B1 a zformulovat doporučení pro lektory, kteří se tuto metodu rozhodnou ve výuce využít.

Vedle uvedení do tématu se první čtyři kapitoly práce věnovaly zejména shrnutí poznatků z prostudované literatury. Postupně jsem se v nich zaměřila na hodnocení jako pedagogický koncept, na vývoj přemýšlení o formativním hodnocení a na dostupná doporučení pro implementaci formativního hodnocení do výuky. Pokusila jsem se také poskytnout co nejucelenější přehled dosavadních výzkumů vzájemného hodnocení ve výuce cizího jazyka.

Pátá kapitola představuje přechod mezi teoretickou a praktickou částí práce. Vysvětluji v ní cíle výzkumu, formuluji výzkumné otázky a hypotézy, popisuji plán a průběh výzkumu a detailněji se zaměřuji také na podobu zkoušek na úrovni B1, z níž provedený experiment vycházel. Zvláštní pozornost přitom věnuji zejména popisu písemných úloh a kritérií hodnocení písemné části, která byla pro přípravu experimentu klíčová, zejména pokud jde o počáteční fázi formulování požadavků, které by texty respondentů měly splňovat.

Získaná data a postup jejich zpracování, stejně jako výsledky výzkumu popisuje šestá kapitola. Přestože provedený experiment nemohl vzhledem k povaze zkoumaného vzorku vést k formulaci obecně platných závěrů, ukázala analýza dat určité tendence, které by ovšem bylo dále nutné ověřit na větším vzorku dat.

Analýza dat naznačuje, že vzájemné hodnocení představuje pro studenty náročnou aktivitu, k níž někteří z nich zaujímají spíše negativní postoj a že zpětnou vazbu od kolegů nehodnotí jako stejně užitečnou v porovnání s pedagogickým hodnocením. V odpovědích na otevřené otázky studenti navíc mnohdy vyjadřovali nejistotu a nedůvěru směrem ke komentářům, které dostali od svých spolužáků. Tento faktor se pravděpodobně promítl také do toho, nakolik byli ochotní přijímat navržené úpravy.

Většina komentářů vzájemného i pedagogického hodnocení se soustředila na návrhy úprav u identifikovaných gramatických a pravopisně-hláskoslovných chyb, pozitivní komentáře se naproti tomu týkaly obvykle textu jako celku. To je do velké míry dáno tím, že jsou právě gramatické a pravopisně-hláskoslovné chyby v získaném vzorku textů nejfrekventovanějšími nedostatky ve vztahu k úrovni B1 a jako takové byly během hodnocení v centru zájmu. Domnívám se, že komentované jevy z dalších kategorií (např. argumentace) by byly více

zastoupeny u komplexnějších textů, které by psali studenti na vyšší jazykové úrovni. Tento předpoklad doporučuji jako námět k dalšímu výzkumu v této oblasti.

Pokud jde o míru povšimnutí chyb, studenti v průměru dosahovali jen přibližně 26% úspěšnosti a při daném přidělení textů jednotliví respondenti dostali v průměru alespoň jeden komentář pouze k 52 % svých chyb. Velké rozdíly v úspěšnosti mezi jednotlivými studenty ovšem naznačují, že vyhledávání chyb představuje velmi specifickou dovednost, která patrně vyžaduje delší nácvik. To je ostatně v souladu s poznatky odborné literatury prezentované v předchozích kapitolách.

Analýza míry správné identifikace nedostatků v textech a správnosti návrhů řešení naznačuje, že studenti většinou nedostatky identifikují správně nebo alespoň částečně správně a k většině z nich nabízejí také správná řešení. Jak už jsem ale nastínila výše, autoři textů často nejsou ochotni tato řešení přijímat. To může ukazovat na jejich nejistotu při rozhodování o správnosti návrhů.

Celkové zhodnocení všech textů provedené další lektorkou sice naznačilo určitý posun v jazykové úrovni některých textů, vzhledem k malému množství dat a nemožnosti vyloučení dalších faktorů ovšem na jeho základě nelze vyvozovat spolehlivé závěry o efektivitě a přínosu zvolených metod.

Domnívám se tedy, že vzájemné hodnocení je možné u studentů připravujících se ke zkoušce B1 využít jako doplňkové cvičení, není ale možné, aby nahradilo pedagogické hodnocení. Aby byla metoda úspěšná jak z hlediska efektivity, tak z hlediska přijetí studenty, je podle mého názoru nutné implementovat ji do výuky postupně v menší míře a poskytnout studentům při prvních pokusech o psaní zpětné vazby jednak více možností nácviku, jednak větší podporu ze strany učitele. S menšími obtížemi pak pravděpodobně bude možné metodu zavádět v homogennějších třídách s pokročilejšími studenty.

Další možnosti výzkumu vzájemného hodnocení ve výuce češtiny jako cizího jazyka tak spatřuji zejména v dlouhodobějším sledování kurzů, v nichž studenti dostanou možnost většího nácviku, který může mít značný vliv na úspěšnost implementace této metody. Nabízí se např. provedení longitudinální studie, v níž by se mohly různým způsobem kombinovat metody jako zúčastněné pozorování, průběžný sběr a analýza textů nebo rozhovory se studenty a lektory.

## 8 Seznam použité literatury

- ALMAHASNEH, Abdallah Mohammad Salem – ABDUL-HAMID, Samsiah. 2019. The Effect of Using Peer Assessment Training on Writing Performance among Arab EFL High School Students in Malaysia. *Arab World English Journal* 10(1). 105–115. (doi:10.24093/awej/vol10no1.10)
- BENDL, Stanislav. 1997. Dotazníkové šetření o subjektivní obtížnosti učitelských činností. *Pedagogika* 47(1). 54–64.
- BENNETT, Randy Elliot. 2011. Formative assessment: a critical review. *Assessment in Education: Principles, Policy & Practice* 18(1). 5–25. (doi:10.1080/0969594X.2010.513678)
- BISCHOFOVÁ, Jana – CONFORTIOVÁ, Helena – CVEJNOVÁ, Jitka – ČADSKÁ, Milada – HOLUB, Jan – LÁNSKÁ, L. – PALKOVÁ, Zdena – TURZÍKOVÁ, Milada. 2001. *Prahová úroveň – čeština jako cizí jazyk*. Vedení projektu Milan Šára. Strasbourg: Council of Europe.
- BLACK, Paul – HARRISON, Christine – LEE, Clara – MARSHALL, Bethan – WILIAM, Dylan. 2003. *Assessment for Learning: putting it into practice*. Maidenhead, U.K.: Open university Press.
- BLACK, Paul – HARRISON, Christine – LEE, Clare – MARSHALL, Bethan – WILIAM, Dylan. 2004. Working inside the Black Box: Assessment for Learning in the Classroom. *Phi Delta Kappan* 86(1). 8–21. (doi:10.1177/003172170408600105)
- BLACK, Paul – MCCORMICK, Robert – JAMES, Mary – PEDDER, David. 2006. Learning How to Learn and Assessment for Learning: a theoretical inquiry. *Research Papers in Education* 21(2). 119–132. (doi:10.1080/02671520600615612)
- BLACK, Paul – WILIAM, Dylan. 2009. Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability* 21(1). 5–31. (doi:10.1007/s11092-008-9068-5)
- BLACK, Paul – WILIAM, Dylan. 2010. Inside the Black Box: Raising Standards through Classroom Assessment. *Phi Delta Kappan* 92(1). 81–90. (doi:10.1177/003172171009200119) [Původní vydání: 1998. *Phi Delta Kappan* 80(2). 139–144, 146–148]
- BLOOM, Benjamin S. 1968. Learning for Mastery. Instruction and Curriculum. *Evaluation Comment* (UCLA-CSEI) 1(2). 1–12.
- BLOOM, Benjamin S. 1971. Mastery Learning. In James H. Block (ed.), *Mastery Learning: Theory and Practice*, 47–63. New York: Holt, Rinehart and Winston.

- BOCCOU KESTŘÁNKOVÁ, Marie – HUDOUSKOVÁ, Andrea – NÝVLTOVÁ, Dana – PEČENÝ, Pavel – ŠVRČINOVÁ, Martina. 2012. *Připravujeme se k certifikované zkoušce z češtiny: úroveň B1 (CCE-B1)*. 1. vyd. Praha: Karolinum.
- BROADFOOT, Patricia – DAUGHERTY, Richard – GARDNER, John – HARLEN, Wynne – JAMES, Mary – STOBART, Gordon. 2002. *Assessment for Learning: 10 Principles. Research-based principles to guide classroom practice* [online]. Assessment Reform Group [cit. 2021-04-19]. Dostupné z: [https://www.researchgate.net/publication/271849158\\_Assessment\\_for\\_Learning\\_10\\_Principles\\_Research-based\\_principles\\_to\\_guide\\_classroom\\_practice\\_Assessment\\_for\\_Learning](https://www.researchgate.net/publication/271849158_Assessment_for_Learning_10_Principles_Research-based_principles_to_guide_classroom_practice_Assessment_for_Learning)
- CVRČEK, Václav. 2017. Precision & recall. In Karlík, Petr – Nekula, Marek – Pleskalová, Jana (eds.), *CzechEncy – Nový encyklopedický slovník češtiny* [online]. Brno: Masarykova univerzita [cit. 2021-05-07]. Dostupné z: [https://www.czechency.org/slovník/PRECISION & RECALL](https://www.czechency.org/slovník/PRECISION&RECALL)
- CVEJNOVÁ, Jitka – ŠORMOVÁ, Kateřina. 2018. Deskriptory pro češtinu jako druhý jazyk. In Suchomelová Lenka (ed.), *Sborník Asociace učitelů češtiny jako cizího jazyka 2017*, 91–102. Akropolis.
- ČECHOVÁ, Marie. 1998. *Komunikační a slohová výchova*. Praha: ISV.
- FALCHIKOV, Nancy – GOLDFINCH, Judy. 2000. Student Peer Assessment in Higher Education: A Meta-Analysis Comparing Peer and Teacher Marks. *Review of Educational Research* 70(3). 287–322. (doi:10.3102/00346543070003287)
- FLOREZ, Maria – SAMMONS, Pamela. 2013. *Assessment for learning: effects and impacts*. Reading: CfBT Education Trust.
- HANRAHAN, Stephanie J. – ISAACS, Geoff. 2001. Assessing Self- and Peer-assessment: The students' views. *Higher Education Research & Development* 20(1). 53–70. (doi:10.1080/07294360123776)
- HRDLIČKA, Milan. 2012. Jazyková chyba a práce s ní v jazykovém vyučování. In Šebesta, Karel – Škodová, Svatava (eds.), *Čeština – cílový jazyk a korpusy*, 89–108. Liberec: TU Liberec.
- HOFSTEDE, Geert. 1983. The Cultural Relativity of Organizational Practices and Theories. *Journal of International Business Studies* 14(2). 75–89. (doi:10.1057/palgrave.jibs.8490867)
- CHANG, C. Yea-huey. 2016. Two Decades of Research in L2 Peer Review. *Journal of Writing Research* 8(1). 81–117. (doi:10.17239/jowr-2016.08.01.03)
- CHENG, Winnie – WARREN, Martin. 2005. Peer assessment of language proficiency. *Language Testing* 22(1). 93–121. (doi:10.1191/0265532205lt298oa)

- INFORMACE O PODOBĚ ZKOUŠKY, 2019. *Certifikovaná zkouška z češtiny pro cizince. Úroveň B1. Informace o podobě zkoušky*. [online]. Ústav jazykové a odborné přípravy Univerzity Karlovy [cit. 2021-04-17]. Dostupné z: [http://ujop.cuni.cz/upload/stories/vtc/CCE/CCE-B1\\_Informace\\_o\\_zkousce.pdf](http://ujop.cuni.cz/upload/stories/vtc/CCE/CCE-B1_Informace_o_zkousce.pdf)
- KALHOUS, Zdeněk. 2002. *Školní didaktika*. Praha: Portál.
- KAUFMAN, Julia H. – SCHUNN, Christian D. 2011. Students' perceptions about peer assessment for writing: their origin and impact on revision work. *Instructional Science* 39(3). 387–406. (doi:10.1007/s11251-010-9133-6)
- LANGAN, A. Mark – SHUKER, David M. – CULLEN, W. Rod – PENNEY, David – Preziosi, Richard F. – WHEATER, C. Philip. 2008. Relationships between student characteristics and self-, peer and tutor evaluations of oral presentations. *Assessment & Evaluation in Higher Education* 33(2). 179–190. (doi:10.1080/02602930701292498)
- LAUFKOVÁ, Veronika. 2018. Formativní hodnocení ve výuce češtiny jako cizího jazyka. In Suchomelová Lenka (ed.), *Sborník Asociace učitelů češtiny jako cizího jazyka 2018*, 81–95. Akropolis.
- LIU, Ngar-Fun – CARLESS, David. 2006. Peer feedback: the learning element of peer assessment. *Teaching in Higher Education* 11(3). 279–290. (doi:10.1080/13562510600680582)
- MANUÁL PRO PŘÍPRAVU K SUBTESTU PSANÍ. 2020. *Certifikovaná zkouška z češtiny pro cizince. Úroveň B1. Manuál pro přípravu k subtestu psaní*. [online]. Ústav jazykové a odborné přípravy Univerzity Karlovy [cit. 2021-04-16]. Dostupné z: [http://ujop.cuni.cz/upload/stories/vtc/CCE/CCE-B1\\_PS\\_Manual\\_pro\\_kandidaty.pdf](http://ujop.cuni.cz/upload/stories/vtc/CCE/CCE-B1_PS_Manual_pro_kandidaty.pdf)
- MODELOVÁ VARIANTA. 2019. *Zkouška z českého jazyka pro cizince. Úroveň B1. Modelová varianta* [online]. Národní pedagogický institut České republiky [cit. 2021-04-17]. Dostupné z: [http://obcanstvi.cestina-pro-cizince.cz/uploads/Modelova\\_varianta\\_zkousky\\_z\\_ceskeho\\_jazyka.pdf](http://obcanstvi.cestina-pro-cizince.cz/uploads/Modelova_varianta_zkousky_z_ceskeho_jazyka.pdf)
- OECD. 2005. *Formative assessment: improving learning in secondary classrooms*. Paris: OECD.
- PASCH, Marvin. 1998. *Od vzdělávacího programu k vyučovací hodině: jak pracovat s kurikulem*. Praha: Portál.
- PRŮCHA, Jan. 2002. *Moderní pedagogika*. 2., přeprac. a aktualiz. vyd. Praha: Portál.
- SERR. 2006. *Společný evropský referenční rámec pro jazyky: jak se učíme jazykům, jak je vyučujeme a jak v jazycích hodnotíme*. 2. české vyd. Olomouc: Univerzita Palackého.
- SLAVÍK, Jan. 1999. *Hodnocení v současné škole: východiska a nové metody pro praxi*. Praha: Portál.



- SMITH, Holly – COOPER, Ali – LANCASTER, Les. 2002. Improving the Quality of Undergraduate Peer Assessment: A Case for Student and Staff Development. *Innovations in Education and Teaching International* 39(1). 71–81. (doi:10.1080/13558000110102904)
- ŠORMOVÁ, Kateřina. 2018. The project of the self-assessment descriptors for Czech as a second language. *Journal of Education Culture and Society* 9(1). 173–180. (doi:10.15503/jecs20181.173.180)
- ŠVRČINOVÁ, Martina – HOUŽVIČKA, Jan (eds.). 2014. *Práce s chybou ve výuce cizích jazyků (včetně češtiny pro cizince)* (Sborník z mezinárodní konference, Poděbrady, 17. – 18. 6. 2014) [online]. Ústav jazykové a odborné přípravy UK [cit. 2021-04-17]. Dostupné z: [http://ujop.cuni.cz/upload/stories/vtc/konference/UJOP\\_2014\\_konference\\_sbornik.pdf](http://ujop.cuni.cz/upload/stories/vtc/konference/UJOP_2014_konference_sbornik.pdf)
- TONUCCI, Francesco. 1991. *Vyučovat nebo naučit?* Praha: Pedagogická fakulta Univerzity Karlovy.
- TORRANCE, Harry. 2012. Formative assessment at the crossroads: conformance, deformative and transformative assessment. *Oxford Review of Education* 38(3). 323–342. (doi:10.1080/03054985.2012.689693)
- VAN ZUNDERT, Marjo – SLUIJSMANS, Dominique – VAN MERRIËNBOER, Jeroen. 2010. Effective peer assessment processes: Research findings and future directions. *Learning and Instruction* 20(4). 270–279. (doi:10.1016/j.learninstruc.2009.08.004)
- VILLAMIL, Olga S. – DE GUERRERO, María C. M. 1998. Assessing the Impact of Peer Revision on L2 Writing. *Applied Linguistics* 19(4). 491–514. (doi:10.1093/applin/19.4.491)
- Vyhláška č. 433/2013 Sb., o prokazování znalosti českého jazyka a českých reálií pro účely udělování státního občanství České republiky.*
- WILLIAM, Dylan. 2010. An integrative summary of the research literature and implications for a new theory of formative assessment. In Andrade, Heidi L. – Cizek, Gregory J. (eds.), *Handbook of formative assessment*, 18–40. New York: Routledge.
- WILLIAMS, Eira. 1992. Student attitudes towards approaches to learning and assessment. *Assessment & Evaluation in Higher Education* 17(1). 45–58. (doi:10.1080/0260293920170105)
- Zákon č. 186/2013 Sb., o státním občanství České republiky.*
- ZKOUŠKA PRO OBČANSTVÍ. 2021. Informace o zkoušce z českého jazyka. *Zkouška pro občanství* [online]. Národní pedagogický institut České republiky [cit. 2021-04-17]. Dostupné z: [https://obcanstvi.cestina-pro-cizince.cz/index.php?p=informace-o-zkousce-2&hl=cs\\_CZ](https://obcanstvi.cestina-pro-cizince.cz/index.php?p=informace-o-zkousce-2&hl=cs_CZ)
- ZKUŠEBNÍ ŘÁD, 2021. *Zkouška z jazyka a reálií pro účely udělování státního občanství České republiky. Zkušební řád* [online]. [cit. 2021-04-17]. Dostupné z: [https://obcanstvi.cestina-pro-cizince.cz/uploads/Zkusebni\\_rad.pdf](https://obcanstvi.cestina-pro-cizince.cz/uploads/Zkusebni_rad.pdf)

## 9 Seznam příloh

Příloha č. 1: Prezentace použita v kurzu

(obsahuje: (i) příkladové texty, (ii) sumarizaci kritérií, (iii) doplňkové instrukce pro studenty, (iv) reflexi první verze dopisu, (v) zadání úloh)

Příloha č. 2: Plakát 10 tipů jak uspět v psaní aneb Pozor na nejčastější chyby

Příloha č. 3: Pretestové úlohy

(obsahuje: (i) zadání, (ii) texty jednotlivých respondentů)

Příloha č. 4: První verze dopisu s pedagogickým hodnocením

Příloha č. 5: Texty ohodnocené vzájemným hodnocením

Příloha č. 6: Druhé verze dopisů

Příloha č. 7: Druhé verze dotazníků

Příloha č. 8: Odpovědi z dotazníků

Příloha č. 9: Obsahová analýza komentářů a míry implementace návrhů úprav ve vzájemném hodnocení

Příloha č. 10: Obsahová analýza komentářů a míry implementace návrhů úprav v pedagogickém hodnocení

Příloha č. 11: Chybová analýza první verze dotazníků

Příloha č. 12: Celkové zhodnocení textů ve vztahu k SERR